

ASYMPTOTICS FOR A CLASS OF SELF-EXCITING POINT PROCESSES

TZU-WEI YANG AND LINGJIONG ZHU

ABSTRACT. In this paper, we study a class of self-exciting point processes. The intensity of the point process has a nonlinear dependence on the past history and time. When a new jump occurs, the intensity increases and we expect more jumps to come. Otherwise, the intensity decays. The model is a marriage between stochasticity and dynamical system. In the short-term, stochasticity plays a major role and in the long-term, dynamical system governs the limiting behavior of the system. We study the law of large numbers, central limit theorem, large deviations and asymptotics for the tail probabilities.

1. INTRODUCTION

Let us consider a class of simple point process N_t with intensity at time t given by

$$(1.1) \quad \lambda_t := \lambda \left(\frac{N_{t-} + \gamma}{t + 1} \right),$$

where $\lambda(\cdot) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is a non-negative function and we also assume that the point process has an empty past history, i.e., $N(-\infty, 0] = 0$. $\gamma \geq 0$ will be called the initial condition and notice that $\lambda_0 = \lambda(\gamma)$. We use $\frac{N_{t-}}{t+1}$ instead of $\frac{N_{t-}}{t}$ to avoid the singularity at $t = 0$. We use N_{t-} instead of N_t in (1.1) to guarantee that the intensity is \mathcal{F}_t -predictable, where \mathcal{F}_t is the natural filtration.

The simple point process N_t by its definition, represents a wide class of self-exciting point processes. When $x \mapsto \lambda(x)$ is an increasing function, the intensity λ_t increases whenever there is a new jump and otherwise it decays. This phenomenon is known as the self-exciting property in the literature. Self-exciting processes have been widely studied in the literature. The self-exciting property makes it ideal to characterize the correlations in some complex systems, including finance. Bacry et al. [1], Bacry et al. [2] studied microstructure noise and Epps effect; Chavez-Demoulin et al. [6] studied value-at-risk; Errais et al. [11] used self-exciting affine point processes to model the credit risk. A Cox-Ingersoll-Ross process with self-exciting jumps is proposed to model the short rate in interest rate models in Zhu [28].

The self-exciting point processes have also been applied to other fields, including seismology, see e.g. Hawkes and Adamopoulos [18], Ogata [23], sociology, see e.g. Crane and Sornette [9] and Blundell et al. [3], and neuroscience, see e.g. Chornoboy et al. [7], Pernice et al. [24], Pernice et al. [25].

Date: 11 December 2014. *Revised:* 11 December 2014.

2000 Mathematics Subject Classification. 60G55, 60F05, 60F10.

Key words and phrases. point processes, self-exciting processes, limit theorems, tail probabilities.

The most popular class of self-exciting point processes is Hawkes process, introduced by Hawkes [16]. It has birth-immigration representation, see Hawkes and Oakes [17] and the limit theorems and Bartlett spectrum have been well studied in the literature, see e.g. Hawkes [16], Bacry et al. [1], Bordenave and Torrisi [4], Zhu [29]. The limit theorems for some variations and extensions of the linear Hawkes processes have been studied in e.g. Karabash and Zhu [19], Zhu [28], Fierro et al. [13], Merhdad and Zhu [22].

Brémaud and Massoulié [5] introduced nonlinear Hawkes process as a generalization of classical Hawkes process. It is a simple point process with intensity

$$(1.2) \quad \lambda_t = \lambda \left(\int_{(-\infty, 0)} h(t-s) N(ds) \right),$$

where $\lambda(\cdot), h(\cdot)$ satisfies certain conditions and many properties of this process, and in particular the limit theorems have been studied recently in Zhu [30], Zhu [31] and Zhu [32]. The name “nonlinear” comes from the nonlinearity of the function $\lambda(\cdot)$. When $\lambda(\cdot)$ is linear, it reduces to the classical Hawkes process. Unlike linear Hawkes process, the nonlinear Hawkes process does not lead to closed-form formulas of the limiting mean, variance in law of large numbers, central limit theorems and the rate function in large deviations.

The model (1.1) proposed in this paper preserves the self-exciting property and nonlinear structure of the nonlinear Hawkes processes while at the same time have more analytical tractability. We also note that if we replace λ_t in (1.1) by $\lambda(N_{t-})$, it becomes the classical pure-birth process, see e.g. Feller [12].

The model (1.1) is time-inhomogeneous Markovian. This can be seen by letting $Y_t := \frac{N_t}{t+1}$ and Y_t satisfies

$$(1.3) \quad dY_t = \frac{\lambda(Y_{t-}) - Y_t}{t+1} dt + \frac{dM_t}{t+1},$$

where $M_t = N_t - \int_0^t \lambda_s ds$ is a martingale. Let us define \bar{Y}_t as the deterministic solution of

$$(1.4) \quad d\bar{Y}_t = \frac{\lambda(\bar{Y}_t) - \bar{Y}_t}{t+1} dt.$$

The limiting behavior of \bar{Y}_t is well understood in dynamical systems. Under the assumptions that there are finitely many fixed points of $x = \lambda(x)$. If we order these fixed points as $x_1 < x_2 < \dots < x_K$, then x_1, x_3, x_5, \dots are stable fixed points and $x_2 < x_4 < \dots$ are unstable fixed points. If \bar{Y}_0 lies on any one of the fixed points, then \bar{Y}_t stays there. Otherwise, \bar{Y}_0 must lie between a stable fixed point and an unstable fixed point and \bar{Y}_t will converge to that neighboring stable fixed point as $t \rightarrow \infty$. This is not true in our model (1.1). For example, if Y_t starts at Y_0 between x_1 and x_2 , it may not end up at x_1 as $t \rightarrow \infty$. That is because there exists a positive probability that the process can jump above x_2 . However, as time t becomes large, the jump size of Y_t that is $\frac{1}{t+1}$ becomes small. Therefore, when time t is small, stochasticity plays a major role in the behavior of (1.1) and when time t is large, the behavior of (1.1) is governed by the dynamical system. Hence our model (1.1) can be seen as a marriage between the dynamical system and stochasticity.

Here are a list of questions we are interested to study.

- If the equation $x \mapsto \lambda(x)$ has a unique fixed point x^* , do we have $\frac{N_t}{t} \rightarrow x^*$?

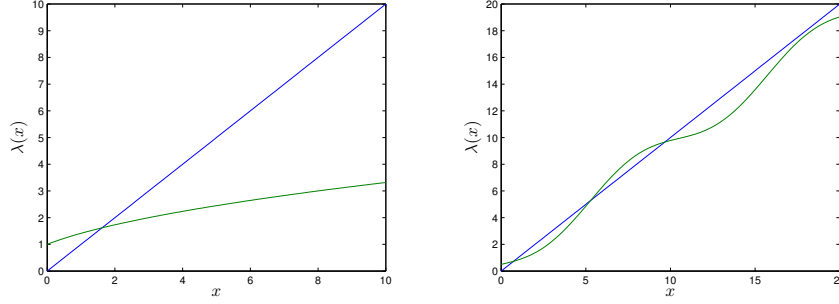


FIGURE 1. On the left hand side, we have the plot of $\lambda(x) = \sqrt{1+x}$ and it has a unique fixed point to the equation $x = \lambda(x)$, $x^* = \frac{1+\sqrt{5}}{2}$. On the right hand side, we have the plot of $\lambda(x) = 0.9x - \sin(0.6x) + 0.5$ and there are three fixed points of the equation $x = \lambda(x)$. Two are stable and the one in between is unstable.

- What if the equation $x = \lambda(x)$ has more than one solution? What should be the limiting set of $\frac{N_t}{t}$ as $t \rightarrow \infty$ then?
- What if $x = \lambda(x)$ has no solutions, what should be the correct scaling for N_t ? And what should be the correct scalings for $\mathbb{E}[N_t]$ and $\text{Var}[N_t]$?
- Can we study the large deviations for $\mathbb{P}(\frac{N_t}{t} \in \cdot)$? And central limit theorems?
- For a fixed time interval $[0, t]$, what is the asymptotics for the tail probabilities $\mathbb{P}(N_t \geq \ell)$ as $\ell \rightarrow \infty$?

We will show that under certain conditions for the model (1.1), the limiting sets of the law of large numbers for $\frac{N_t}{t}$ is the set of stable fixed points of $x = \lambda(x)$. When the equation $x = \lambda(x)$ has a unique fixed point, the limit is therefore the unique fixed point. It gets more interesting when there are more than one fixed point. Second-order properties will also be studied, including the variance and covariance structure. A sample-path large deviation principle will be derived and hence the large deviations for $\mathbb{P}(N_t/t \in \cdot)$ as well.

Figure 3 illustrates that when there is a unique fixed point of $x = \lambda(x)$, as time $t \rightarrow \infty$, $\frac{N_t}{t}$ converges to this unique fixed point. When there are more than one fixed point, Figure 4 illustrates that as time $t \rightarrow \infty$, $\frac{N_t}{t}$ will converge to the set of all the stable fixed points of $x = \lambda(x)$. Let us say in Figure 4, the two stable fixed points are $x_1 < x_2$. Let

$$(1.5) \quad p_1(x) := \mathbb{P}_{N_0=x} \left(\lim_{t \rightarrow \infty} \frac{N_t}{t} = x_1 \right), \quad p_2(x) := \mathbb{P}_{N_0=x} \left(\lim_{t \rightarrow \infty} \frac{N_t}{t} = x_2 \right).$$

Then, for any initial condition $N_0 = x$, $p_1(x) + p_2(x) = 1$. Intuitively, it is clear that when $N_0 = x$ is closer to x_1 than x_2 is in between x_1 and x_2 , there should be a higher probability for the limit $\lim_{t \rightarrow \infty} \frac{N_t}{t}$ to end up at x_1 . If the starting point is lower than x_1 , it is also more likely for the limit to end up at x_1 and so on and so forth. We can therefore use the same $\lambda(x)$ as in Figure 4 and make a plot of p_1 and p_2 as a function of the initial starting point $N_0 = x$. From Figure 5, it turns out

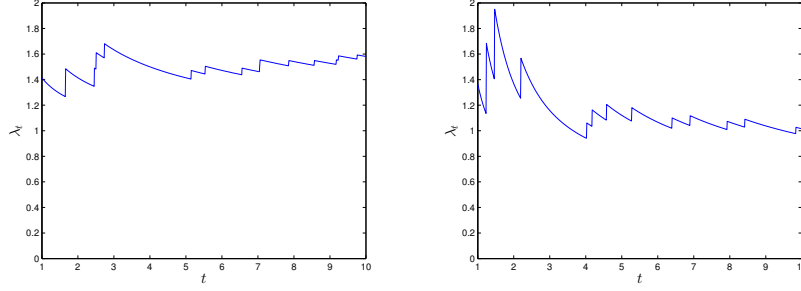


FIGURE 2. On the left hand side, we have the plot of λ_t for $\lambda(x) = \sqrt{1+x}$. On the right hand side, we have the plot of λ_t for $\lambda(x) = 0.9x - \sin(0.6x) + 0.5$. We zoom into the time interval $[1, 10]$ to see the local self-exciting behavior of the model.

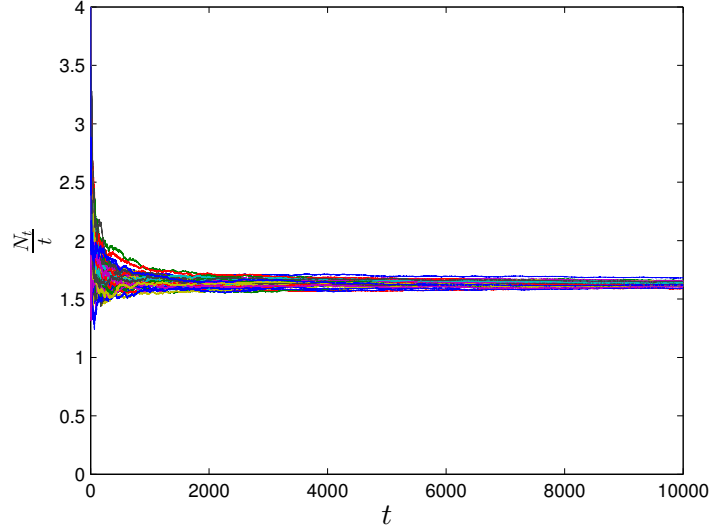


FIGURE 3. $\lambda(x) = \sqrt{1+x}$ and it has a unique fixed point to the equation $x = \lambda(x)$, $x^* = \frac{1+\sqrt{5}}{2}$. The initial condition is assumed to be $N_0 = 5$. As time $t \rightarrow \infty$, $\frac{N_t}{t}$ converges to this unique fixed point.

that $p_1(x)$ is monotonically decreasing in x and $p_2(x) = 1 - p_1(x)$ is monotonically increasing in x .

The paper is organized as follows. In Section 2 we will state all the main results. In particular, we will discuss the law of large numbers in Section 2.1, large deviations in Section 2.2, time asymptotics for different regimes in Section 2.3, asymptotics for high initial values in Section 2.4 and marginal and tail probabilities in Section

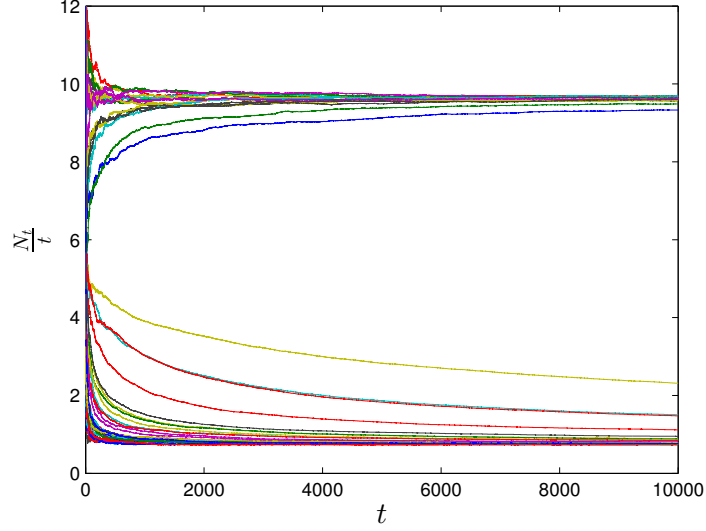


FIGURE 4. $\lambda(x) = 0.9x - \sin(0.6x) + 0.5$ and there are three fixed points of the equation $x = \lambda(x)$. The initial condition is assumed to be $N_0 = 5$. Two are stable and the one in between is unstable. As time $t \rightarrow \infty$, $\frac{N_t}{t}$ will converge to either one of the stable fixed points.

2.5. The proofs will be given in Section 3. Finally some open problems will be suggested in Section 4.

2. MAIN RESULTS

Throughout the paper, we assume the following conditions hold.

- $\lambda : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is an increasing and continuously differentiable function.
- $x \mapsto \lambda(x)$ has finitely many fixed points, i.e., the equation $x = \lambda(x)$ has finitely many solutions. The fixed points are either strictly stable or strictly unstable, i.e., if x^* is a fixed point, then either $\lambda'(x^*) < 1$ or $\lambda'(x^*) > 1$.

2.1. Law of Large Numbers. Assume that $\lambda(z) \leq \beta + \alpha z$ for some $0 < \alpha < 1$ and $\beta > 0$ and $\lambda(x) = x$ has a unique solution x^* . Then, by Proposition 14, $\sup_{t \geq 0} \frac{\mathbb{E}[N_t]}{t+1} < \infty$. Thus $\frac{N_t}{t+1}$ is tight. Heuristically, if we have $\frac{N_t}{t} \rightarrow x$ a.s. as $t \rightarrow \infty$. Then, we have as $t \rightarrow \infty$, a.s.,

$$(2.1) \quad \frac{N_t}{t} \rightarrow x, \quad \frac{1}{t} \int_0^t \lambda\left(\frac{N_{s-}}{s+1}\right) ds \rightarrow \lambda(x).$$

Moreover, $M_t = N_t - \int_0^t \lambda_s ds$ is a martingale and

$$(2.2) \quad \mathbb{E}[(M_t)^4] \leq C\mathbb{E}[N_t^2] = O(t^2)$$

by Burkholder-Davis-Gundy inequality and Proposition 14. Thus, $\frac{M_t}{t} \rightarrow 0$ in a.s. by Borel-Cantelli lemma. Since $\lambda(x) = x$ has a unique solution x^* , we conclude that $\frac{N_t}{t} \rightarrow x^*$ a.s.

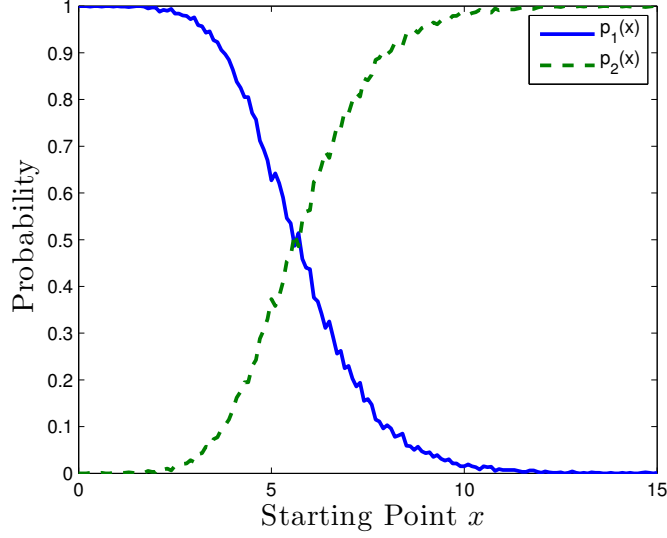


FIGURE 5. Plot of the probabilities that $\frac{N_t}{t}$ will converge to $x_1 < x_2$, the stable fixed point of $x = \lambda(x) = 0.9x - \sin(0.6x) + 0.5$ as a function of the starting positions.

Theorem 1. Assume that $\lambda(z)$ is increasing, α -Lipschitz with $0 < \alpha < 1$ and x^* is the unique solution to the equation $x = \lambda(x)$ and $x^* = \infty$ if the solution does not exist, then

$$(2.3) \quad \frac{N_t}{t} \rightarrow x^*,$$

in probability as $t \rightarrow \infty$. If we further assume that $\lambda(\cdot) \leq C_0 < \infty$ for some universal constant C_0 , then we have the almost sure convergence.

Remark 2. In Theorem 1, if $\lambda(z) \leq \beta + \alpha z$ for some $0 < \alpha < 1$, $\beta > 0$ and $\lambda(z)$ is continuous, then the equation $\lambda(x) = x$ must have at least one solution. In the case that 0 is the only solution to $\lambda(x) = x$, we have $\frac{N_t}{t} \rightarrow 0$ in probability as $t \rightarrow \infty$. On the other hand, if $\lambda(z)$ is continuous and the equation $\lambda(x) = x$ has no solution, then we must have $\lambda(z) > z$ for any $z \geq 0$. Hence, we have $\lambda(z) \geq (1 - \epsilon)z + \delta$ for some $\delta, \epsilon > 0$ sufficiently small. Note that $(1 - \epsilon)z + \delta = z$ if and only if $z = \frac{\delta}{\epsilon}$. Choose $\delta \gg \epsilon$ and by Theorem 1, we conclude that $\frac{N_t}{t} \rightarrow \infty$ in probability as $t \rightarrow \infty$ if $\lambda(z)$ is continuous and the equation $\lambda(x) = x$ has no solution.

In Theorem 1, we proved that $\frac{N_t}{t}$ converges to the unique fixed point of $x = \lambda(x)$ in probability under the α -Lipschitz condition for $\lambda(\cdot)$ for some $0 < \alpha < 1$ and proved that the convergence is a.s. convergence under a stronger condition. Next, we compare the underlying stochastic process $Y_t := \frac{N_t}{t+1}$ to its deterministic counterpart \bar{Y}_t where \bar{Y}_t is the deterministic solution of

$$(2.4) \quad d\bar{Y}_t = \frac{\lambda(\bar{Y}_t) - \bar{Y}_t}{t+1} dt,$$

whose asymptotic behavior is entirely governed by the dynamical system and prove a law of large numbers in the $L^2(\mathbb{P})$ norm. As a by-product, we also get the convergence rate of the underlying stochastic process to its deterministic counterpart.

Theorem 3. *Assume that $\lambda(\cdot)$ is α -Lipschitz for some $0 < \alpha < 1$, then $Y_t = \frac{N_t}{t+1}$ converges to the unique fixed point of $x = \lambda(x)$ as $t \rightarrow \infty$ in $L^2(\mathbb{P})$ norm. Moreover, as $t \rightarrow \infty$,*

$$(2.5) \quad \mathbb{E}[(Y_t - \bar{Y}_t)^2] = \begin{cases} O(\frac{1}{t}) & \text{if } 0 < \alpha < \frac{1}{2} \\ O(\frac{1}{t^{2(1-\alpha)}}) & \text{if } \frac{1}{2} < \alpha < 1 \\ O(\frac{\log(t)}{t}) & \text{if } \alpha = \frac{1}{2} \end{cases}.$$

Theorem 4. *Assume that $x \mapsto \lambda(x)$ is continuous and increasing. For any interval $I = [a, b]$ not containing any fixed point of the equation $x = \lambda(x)$, we have*

$$(2.6) \quad \lim_{t \rightarrow \infty} \mathbb{P}\left(\frac{N_t}{t} \in I\right) = 0.$$

Theorem 5. *Let x^* be any stable fixed point of $\lambda(x)$: $x^* = \lambda(x^*)$ and $0 < \lambda'(x^*) < 1$, the probability that $\frac{N_t}{t} \rightarrow x^*$ is non-zero.*

Theorem 6. *Let x^* be any unstable fixed point of $\lambda(x)$: $x^* = \lambda(x^*)$ and $\lambda'(x^*) > 1$, the probability that $\frac{N_t}{t} \rightarrow x^*$ is zero.*

2.2. Large Deviations. Before we proceed, recall that a sequence $(P_n)_{n \in \mathbb{N}}$ of probability measures on a topological space \mathbb{X} satisfies the large deviation principle with rate function $\mathcal{I} : \mathbb{X} \rightarrow \mathbb{R}$ if \mathcal{I} is non-negative, lower semicontinuous and for any measurable set A , we have

$$(2.7) \quad -\inf_{x \in A^\circ} \mathcal{I}(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log P_n(A) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log P_n(A) \leq -\inf_{x \in \bar{A}} \mathcal{I}(x).$$

Here, A° is the interior of A and \bar{A} is its closure. We refer to the books by Dembo and Zeitouni [10] and Varadhan [27] and the survey paper by Varadhan [26] for general background of the theory and the applications of large deviations.

Theorem 7. *Assume that $\lambda(\cdot) \leq C_0 < \infty$ and $\lambda(\cdot)$ is γ -Lipschitz for some $0 < \gamma < \infty$. $\mathbb{P}(\frac{N_T}{T} \in \cdot)$ satisfies a sample path large deviations on $D[0, 1]$ equipped with uniform topology with the rate function*

$$(2.8) \quad I(f) = \int_0^1 \log \left(\frac{f'(\alpha)}{\lambda(f(\alpha)/\alpha)} \right) f'(\alpha) d\alpha - \int_0^1 [f'(\alpha) - \lambda(f(\alpha)/\alpha)] d\alpha.$$

By contraction principle, we get the following scalar large deviation principle.

Corollary 8. $\mathbb{P}(N_t/t \in \cdot)$ satisfies a large deviation principle with rate function

$$(2.9) \quad I(x) := \inf_{f \in \mathcal{A}_0^+[0, 1], f(1)=x} \int_0^1 \log \left(\frac{f'(\alpha)}{\lambda(f(\alpha)/\alpha)} \right) f'(\alpha) d\alpha - \int_0^1 [f'(\alpha) - \lambda(f(\alpha)/\alpha)] d\alpha.$$

Remark 9. *It would be interesting to see if one can relax the assumption $\lambda(\cdot) \leq C_0 < \infty$. This may not be easy or even possible. For instance, if $\lambda(z) \geq \alpha z$ for some $\alpha > 0$, then for any $\theta > 0$, by (2.36), we have*

$$(2.10) \quad \mathbb{E}[e^{\theta N_t}] \geq \begin{cases} \left(\frac{\frac{1}{(t+1)^\alpha}}{1 - \frac{1}{(t+1)^\alpha} e^\theta} \right)^\gamma & \text{if } t < (1 - e^{-\theta})^{-\frac{1}{\alpha}} - 1 \\ \infty & \text{if } t \geq (1 - e^{-\theta})^{-\frac{1}{\alpha}} - 1 \end{cases}.$$

Thus for any $\theta > 0$, $\mathbb{E}[e^{\theta N_t}] = \infty$ for sufficiently large t .

Remark 10. Let us define

$$(2.11) \quad L(\alpha, f(\alpha), f'(\alpha)) = \log \left(\frac{f'(\alpha)}{\lambda(f(\alpha)/\alpha)} \right) f'(\alpha) - [f'(\alpha) - \lambda(f(\alpha)/\alpha)].$$

Thus, we are interested to optimize $\int_0^1 L(\alpha, f(\alpha), f'(\alpha)) d\alpha$ subject to the constraints $f(0) = 0$ and $f(1) = x$. We can write down the Euler-Lagrange equation

$$(2.12) \quad \begin{aligned} 0 &= \frac{\partial L}{\partial f} - \frac{d}{d\alpha} \frac{\partial L}{\partial f'} \\ &= -\frac{\lambda'(f(\alpha)/\alpha)}{\alpha \lambda(f(\alpha)/\alpha)} + \frac{\lambda'(f(\alpha)/\alpha)}{\alpha} - \frac{d}{d\alpha} \left[\log(f'(\alpha)) - f'(\alpha) \log \lambda \left(\frac{f(\alpha)}{\alpha} \right) \right]. \end{aligned}$$

Remark 11. If $x = \lambda(x)$, by letting $f(\alpha) = \alpha x$, $0 \leq \alpha \leq 1$, we can easily check that $I(x) = 0$.

Proposition 12. The converse of Remark 11 is also true. In other words, if $I(x) = 0$, then x must be a fixed point of $\lambda(x)$ and the minimizer is $f(\alpha) = \alpha x$.

Remark 13. Let x^* be any unstable fixed point of $x \mapsto \lambda(x)$ and C be any sufficiently small neighborhood containing x^* . From Theorem 4 and Theorem 6, we have $\mathbb{P}(\frac{N_t}{t} \in C) \rightarrow 0$ as $t \rightarrow \infty$. On the other hand, from Proposition 12, we have $I(x^*) = 0$, which implies that $\mathbb{P}(\frac{N_t}{t} \in C)$ has subexponential decay in time as $t \rightarrow \infty$. This is consistent with our simulations results which illustrate that when there is a configuration in which $\frac{N_t}{t}$ is in a small neighborhood of x^* , it takes a very long time for the process to exit the neighborhood.

2.3. Time Asymptotics in Different Regimes. When $\lambda(\cdot)$ is α -Lipschitz with $0 < \alpha < 1$, we know that there is a unique fixed point to $x = \lambda(x)$ and $\frac{N_t}{t}$ converges to this unique fixed point as $t \rightarrow \infty$. In general, $x = \lambda(x)$ may not have any fixed points. If that's the case, then what should be the correct scaling for N_t , $\mathbb{E}[N_t]$, $\text{Var}[N_t]$ etc.? In this section, we study the different time asymptotics for different regimes.

Proposition 14. Assume $\lambda(z) = \beta + \alpha z$, $\alpha, \beta > 0$ and $\alpha \neq 1$.

(i) The expectation is given by

$$(2.13) \quad \mathbb{E}[N_t] = (t+1) \frac{\beta}{1-\alpha} [1 - (t+1)^{\alpha-1}],$$

and thus for $\alpha < 1$

$$(2.14) \quad \lim_{t \rightarrow \infty} \frac{\mathbb{E}[N_t]}{t} = \frac{\beta}{1-\alpha},$$

and for $\alpha > 1$

$$(2.15) \quad \lim_{t \rightarrow \infty} \frac{\mathbb{E}[N_t]}{t^\alpha} = \frac{\beta}{\alpha-1}.$$

(ii) The variance is given by

$$(2.16) \quad \text{Var}[N_t] = (t+1)^2 \left[\frac{\beta}{(1-2\alpha)(1-\alpha)} + \frac{\beta}{(t+1)^{2-\alpha}} + \frac{-\frac{2\beta}{1-2\alpha}}{(t+1)^{2(1-\alpha)}} \right].$$

For $0 < \alpha < \frac{1}{2}$, we have

$$(2.17) \quad \lim_{t \rightarrow \infty} \frac{\text{Var}[N_t]}{t} = \frac{\beta}{(1-2\alpha)(1-\alpha)}.$$

For $\alpha > \frac{1}{2}$, we have

$$(2.18) \quad \lim_{t \rightarrow \infty} \frac{\text{Var}[N_t]}{t^{2\alpha}} = \frac{2\beta}{2\alpha-1}.$$

Theorem 15. Assume that $\lambda(z) = \beta + \alpha z$, $\alpha, \beta > 0$ and $\alpha < \frac{1}{2}$. Then, we have the central limit theorem

$$(2.19) \quad \frac{N_t - \frac{\beta}{1-\alpha}t}{\sqrt{t}} \rightarrow N\left(0, \frac{\beta}{(1-2\alpha)(1-\alpha)}\right),$$

in distribution as $t \rightarrow \infty$.

Proposition 16. For $\lambda(z) = \beta + \frac{1}{2}z$, $\beta > 0$,

$$(2.20) \quad \text{Var}[N_t] = (t+1)^2 \left[\frac{2\beta[\log(t+1)-1]}{t+1} + \frac{2\beta}{(t+1)^{3/2}} \right],$$

and

$$(2.21) \quad \lim_{t \rightarrow \infty} \frac{\text{Var}[N_t]}{t \log t} = 2\beta.$$

Proposition 17. For $\lambda(z) = \beta + z$, $\beta > 0$,

$$(2.22) \quad \mathbb{E}[N_t] = \beta(t+1) \log(t+1),$$

and

$$(2.23) \quad \text{Var}[N_t] = -\beta(t+1) \log(t+1) + 2\beta t(t+1)$$

and

$$(2.24) \quad \lim_{t \rightarrow \infty} \frac{\text{Var}[N_t]}{t^2} = 2\beta.$$

Corollary 18. Let $\lambda(z) = \beta + \alpha z$, $\beta, \alpha > 0$, then

- (i) For $\alpha < 1$, $\frac{N_t}{t} \rightarrow \frac{\beta}{1-\alpha}$ in probability.
- (ii) For $\alpha > 1$, $\frac{N_t}{t^\alpha} \rightarrow \frac{\beta}{\alpha-1}$ in probability.
- (iii) For $\alpha = 1$, $\frac{N_t}{t \log t} \rightarrow \beta$ in probability.

We can also compute the covariance structure explicitly when $\lambda(\cdot)$ is linear.

Proposition 19. Assume $\lambda(z) = \beta + \alpha z$, $\alpha, \beta > 0$ and $\alpha \notin \{\frac{1}{2}, 1\}$, for any $t > s$,

$$(2.25) \quad \text{Cov}[N_t, N_s] = (t+1)^\alpha \left[\frac{\beta}{(1-2\alpha)(1-\alpha)} (s+1)^{1-\alpha} + \frac{\beta}{1-\alpha} - \frac{2\beta}{1-2\alpha} (s+1)^\alpha \right].$$

For $\alpha = \frac{1}{2}$,

$$(2.26) \quad \text{Cov}[N_t, N_s] = 2\beta [-(t+1)^{\frac{1}{2}}(s+1)^{\frac{1}{2}} + (t+1)^{\frac{1}{2}} + (t+1)^{\frac{1}{2}}(s+1)^{\frac{1}{2}} \log(s+1)].$$

Proposition 20. Assume $\lambda(z) = \beta + z$, $\beta > 0$, for any $t > s$,

$$(2.27) \quad \text{Cov}[N_t, N_s] = -\beta(t+1) \log(s+1) + 2\beta s(t+1).$$

We have seen that if $\lambda(z) = \alpha z$, $0 < \alpha < 1$, we have $\lim_{t \rightarrow \infty} \frac{N_t}{t} = 0$. A natural question to ask is under this regime, what should be the correct scaling for N_t as time t goes to ∞ . Since $\lambda(0) = 0$, we need to start the process at some positive inditial condition $\gamma > 0$.

Proposition 21. *Let the intensity at time t be*

$$(2.28) \quad \lambda_t = \frac{\alpha(N_{t-} + \gamma)}{t + 1}, \quad \alpha, \gamma > 0.$$

Then, we have $\mathbb{E}[N_t] = \gamma[(t + 1)^\alpha - 1]$ and

$$(2.29) \quad \text{Var}[N_t] = \gamma[(t + 1)^{2\alpha} - (t + 1)^\alpha].$$

In particular, $\lim_{t \rightarrow \infty} \frac{\mathbb{E}[N_t]}{t^\alpha} = \gamma$, and $\lim_{t \rightarrow \infty} \frac{\text{Var}[N_t]}{t^{2\alpha}} = \gamma$. Also, for any $t > s$,

$$(2.30) \quad \text{Cov}[N_t, N_s] = [(s + 1)^\alpha - 1] \left[\gamma(t + 1)^\alpha + (\gamma - \gamma^2) \left[\frac{(t + 1)^\alpha}{(s + 1)^\alpha} - 1 \right] \right].$$

Moreover, as $t \rightarrow \infty$,

$$(2.31) \quad \frac{N_t}{t^\alpha} \rightarrow \chi(\gamma),$$

a.s. and in $L^2(\mathbb{P})$ where $\chi(\gamma)$ is a random variable with gamma distribution with parameters γ (shape) and 1 (scale).

We end this section with a criterion on whether the point process N_t can be explosive or not. Essentially, when $\lambda(\cdot)$ is super linear, it gives the explosive regime.

Proposition 22. *Assume that $\int_0^\infty \frac{1}{\lambda(z)} dz < \infty$. Then, the point process is explosive. More precisely, $0 < \mathbb{P}(\tau < \infty) < 1$, where $\tau := \inf\{t > 0 : N_t = \infty\}$.*

2.4. High Initial Value. One can also study the asymptotics for high initial value $\gamma \rightarrow \infty$. In the classical birth-death process, that corresponds to high initial population size. The asymptotics results for high initial values can be interesting and useful. For example, they are useful in the models of cancer dyanmics, see e.g. Foo and Leder [14], Foo et al. [15].

Proposition 23. *Assume that $\lim_{z \rightarrow \infty} \frac{\lambda(z)}{z} = \alpha$. Then,*

$$(2.32) \quad \sup_{0 \leq s \leq t} \left| \frac{N_s}{\gamma} - [(s + 1)^\alpha - 1] \right| \rightarrow 0,$$

in probability as $\gamma \rightarrow \infty$.

We can also study the case when $\lambda(\cdot)$ is sublinear.

Proposition 24. *Assume that $\lim_{z \rightarrow \infty} \frac{\lambda(z)}{z^\beta} = \alpha$, where $\alpha > 0$ and $0 < \beta < 1$. Then,*

$$(2.33) \quad \sup_{0 \leq s \leq t} \left| \frac{N_s}{\gamma^\beta} - \frac{\alpha}{1 - \beta} [(s + 1)^{1-\beta} - 1] \right| \rightarrow 0,$$

in probability as $\gamma \rightarrow \infty$.

2.5. Marginal and Tail Probabilities. In this section, we are interested to study the marginal probabilities $\mathbb{P}(N_t = k)$ for a given $k \in \mathbb{N} \cup \{0\}$ and the asymptotics for the tail probabilities $\mathbb{P}(N_t \geq \ell)$ for large ℓ . We assume that the initial condition is given by $\lambda_0 = \lambda(\gamma)$, where $\gamma \in \mathbb{R}^+$.

Theorem 25. *For any $k \in \mathbb{N} \cup \{0\}$,*

(i)
(2.34)

$$\mathbb{P}(N_t = k) = \int \cdots \int_{0 < t_1 < t_2 < \cdots < t_k < t} \prod_{j=1}^k \lambda \left(\frac{\gamma + j - 1}{t_j + 1} \right) \cdot e^{-\int_0^{t_1} \lambda(\frac{\gamma}{s+1}) ds - \int_{t_1}^{t_2} \lambda(\frac{\gamma+1}{s+1}) ds - \cdots - \int_{t_k}^t \lambda(\frac{\gamma+k}{s+1}) ds} dt_1 dt_2 \cdots dt_k.$$

(ii) *In particular, the void probability is given by*

$$(2.35) \quad \mathbb{P}(N_t = 0) = e^{-\int_0^t \lambda(\frac{\gamma}{s+1}) ds}.$$

(iii) *When $\lambda(z) = \alpha z$, N_t follows a negative binomial distribution,*

$$(2.36) \quad \mathbb{P}(N_t = k) = \binom{k + \gamma - 1}{k} \left(1 - \frac{1}{(t+1)^\alpha} \right)^k \left(\frac{1}{(t+1)^\alpha} \right)^\gamma.$$

For any $t > s > 0$ and $k, m \in \mathbb{N} \cup \{0\}$, we have the conditional probability

(2.37)

$$\mathbb{P}(N_t = k + m | N_s = m) = \binom{k + m + \gamma - 1}{k} \left(1 - \left(\frac{s+1}{t+1} \right)^\alpha \right)^k \left(\left(\frac{s+1}{t+1} \right)^\alpha \right)^{\gamma+m}.$$

For a standard Poisson process N_t with constant intensity λ , the tail probability $\mathbb{P}(N_t \geq \ell)$ has the asymptotics $\lim_{\ell \rightarrow \infty} \frac{1}{\ell \log \ell} \log \mathbb{P}(N_t \geq \ell) = -1$. What is the asymptotics for the tail probabilities in our model? In the next result, we will show that if $\lambda(\cdot)$ is asymptotically linear, then unlike the standard Poisson process, we have exponential tails for $\mathbb{P}(N_t \geq \ell)$ as ℓ goes to infinity.

Theorem 26. *Assume that $\lim_{z \rightarrow \infty} \frac{\lambda(z)}{z} = \alpha \in (0, \infty)$. Then, for any fixed $t > 0$,*

$$(2.38) \quad \lim_{\ell \rightarrow \infty} \frac{1}{\ell} \log \mathbb{P}(N_t \geq \ell) = \log \left(1 - \frac{1}{(t+1)^\alpha} \right).$$

We have already studied the tail probabilities for $\mathbb{P}(N_t \geq \ell)$ when $\lambda(\cdot)$ is asymptotically linear in Theorem 26. One can also study the case when $\lim_{z \rightarrow \infty} \frac{\lambda(z)}{z^\beta} = \alpha$, for some $\alpha, \beta > 0$. When $\beta > 1$, $\lambda(z)$ grows super-linearly and there is a positive probability of explosion. Therefore in this case, $\mathbb{P}(N_t \geq \ell)$ does not vanish to zero as $\ell \rightarrow \infty$. When $\beta < 1$, $\lambda(z)$ grows sub-linearly and $\mathbb{P}(N_t \geq \ell)$ does vanish to zero as $\ell \rightarrow \infty$. The asymptotics of the tail probabilities are studied as follows.

Theorem 27. *Assume that $\lim_{z \rightarrow \infty} \frac{\lambda(z)}{z^\beta} = \alpha$, for some $\alpha, \beta > 0$ and $\beta < 1$. Then,*

$$(2.39) \quad \lim_{\ell \rightarrow \infty} \frac{1}{\ell \log \ell} \log \mathbb{P}(N_t \geq \ell) = -(1 - \beta).$$

Remark 28. *For any $\lambda(z)$ that grows slower than any polynomial growth, the tail is the same as the Poisson tail from Theorem 27. For example, for $\lambda(z)$ uniformly bounded, $\lambda(z) = [\log(1+z)]^\beta$, $\beta > 0$, they all give the Poisson tail $\lim_{\ell \rightarrow \infty} \frac{1}{\ell \log \ell} \log \mathbb{P}(N_t \geq \ell) = -1$.*

3. PROOFS

3.1. Proofs of Results in Section 2.1.

Proof of Theorem 1. Let us use Poisson embedding. Let $N^{(0)}$ be the Poisson process with intensity $\lambda(0)$. Conditional on $N^{(0)}$, let $N^{(1)}$ be the inhomogeneous Poisson process with intensity

$$(3.1) \quad \lambda \left(\frac{N_{t-}^{(0)}}{t+1} \right) - \lambda(0),$$

at time t . Inductively, conditional on $N^{(0)}, N^{(1)}, \dots, N^{(k)}$, $N^{(k+1)}$ is an inhomogeneous Poisson process with intensity

$$(3.2) \quad \lambda \left(\frac{N_{t-}^{(0)} + N_{t-}^{(1)} + \dots + N_{t-}^{(k)}}{t+1} \right) - \lambda \left(\frac{N_{t-}^{(0)} + N_{t-}^{(1)} + \dots + N_{t-}^{(k-1)}}{t+1} \right),$$

at time t . Therefore, we can compute that $\mathbb{E}[N_t^{(0)}] = \lambda(0)t$,

$$(3.3) \quad \begin{aligned} \frac{\mathbb{E}[N_t^{(1)}]}{t} &= \frac{1}{t} \mathbb{E} \left[\int_0^t \lambda \left(\frac{N_{s-}^{(0)}}{s+1} \right) - \lambda(0) ds \right] \\ &\leq \alpha \frac{1}{t} \int_0^t \frac{\mathbb{E}[N_{s-}^{(0)}]}{s+1} ds \\ &\leq \alpha \lambda(0). \end{aligned}$$

and inductively,

$$(3.4) \quad \frac{1}{t} \mathbb{E}[N_t^{(k)}] \leq \alpha^k \lambda(0), \quad k \in \mathbb{N}.$$

Hence, $\mathbb{E}[\sum_{k=0}^{\infty} N_t^{(k)}] \leq \sum_{k=0}^{\infty} \alpha^k \lambda(0)t < \infty$ since $0 < \alpha < 1$ and $N_t = \sum_{k=0}^{\infty} N_t^{(k)}$ is well defined a.s. Moreover, the compensator of N_t is

$$(3.5) \quad \sum_{k=0}^{\infty} \int_0^t \left[\lambda \left(\frac{\sum_{j=0}^k N_{s-}^{(j)}}{s+1} \right) - \lambda \left(\frac{\sum_{j=0}^{k-1} N_{s-}^{(j)}}{s+1} \right) \right] ds = \int_0^t \lambda \left(\frac{N_{s-}}{s+1} \right) ds$$

and hence N_t is the self-exciting process we are interested to study. By the law of large numbers for Poisson processes,

$$(3.6) \quad \frac{N_t^{(0)}}{t} \rightarrow \lambda(0),$$

a.s. as $t \rightarrow \infty$. We use induction and assume that $\frac{N_t^{(j)}}{t} \rightarrow \lambda^{(j+1)}(0) - \lambda^{(j)}(0)$ a.s. as $t \rightarrow \infty$ for any $j = 0, 1, \dots, k-1$. Then, the compensator $\Lambda_t^{(k)}$ of $N_t^{(k)}$ satisfies

$$(3.7) \quad \begin{aligned} \frac{1}{t} \int_0^t \lambda \left(\frac{N_{s-}^{(0)} + N_{s-}^{(1)} + \dots + N_{s-}^{(k-1)}}{s+1} \right) - \lambda \left(\frac{N_{s-}^{(0)} + N_{s-}^{(1)} + \dots + N_{s-}^{(k-2)}}{s+1} \right) ds \\ \rightarrow \lambda^{(k+1)}(0) - \lambda^{(k)}(0), \end{aligned}$$

a.s. as $t \rightarrow \infty$. On the other hand, $M_t^{(k)} := N_t^{(k)} - \Lambda_t^{(k)}$ is a martingale and

$$(3.8) \quad \frac{M_t^{(k)}}{t} \rightarrow 0,$$

a.s. as $t \rightarrow \infty$. Hence, we proved that

$$(3.9) \quad \frac{1}{t} \sum_{j=0}^k N_t^{(j)} \rightarrow \lambda^{(k+1)}(0),$$

a.s. as $t \rightarrow \infty$. As $k \rightarrow \infty$, $\lambda^{(k+1)}(0) \rightarrow x^*$. For any $\epsilon > 0$, there exists $K \in \mathbb{N}$ so that for any $k \geq K$, $|\lambda^{(k+1)}(0) - x^*| < \frac{\epsilon}{4}$. Thus,

$$(3.10) \quad \begin{aligned} & \limsup_{t \rightarrow \infty} \mathbb{P} \left(\left| \sum_{k=0}^{\infty} \frac{N_t^{(k)}}{t} - x^* \right| \geq \epsilon \right) \\ & \leq \limsup_{t \rightarrow \infty} \mathbb{P} \left(\left| \sum_{k=0}^K \frac{N_t^{(k)}}{t} - \lambda^{(K+1)}(0) \right| \geq \frac{\epsilon}{2} \right) \\ & \quad + \limsup_{t \rightarrow \infty} \mathbb{P} \left(\left| \sum_{k=K+1}^{\infty} \frac{N_t^{(k)}}{t} - (x^* - \lambda^{(K+1)}(0)) \right| \geq \frac{\epsilon}{2} \right) \\ & \leq \limsup_{t \rightarrow \infty} \mathbb{P} \left(\sum_{k=K+1}^{\infty} \frac{N_t^{(k)}}{t} \geq \frac{\epsilon}{4} \right) \\ & \leq \frac{\lambda(0) \sum_{k=K+1}^{\infty} \alpha^k}{(\epsilon/4)}. \end{aligned}$$

Since it holds for any $K \in \mathbb{N}$, we get the desired result by letting $K \rightarrow \infty$.

Finally, if we further assume that $\lambda(\cdot) \leq C_0 < \infty$ for some universal constant c_0 and C_0 , then by the large deviations results in Theorem 7 and Proposition 12, we get the almost sure convergence. \square

Proof of Theorem 3. Let us define $Y_t = \frac{N_t}{t+1}$. It is easy to check that

$$(3.11) \quad dY_t = \frac{\lambda(Y_{t-}) - Y_t}{t+1} dt + \frac{dM_t}{t+1},$$

where $M_t = N_t - \int_0^t \lambda_s ds$ is a martingale. Let us define \bar{Y}_t as the deterministic solution of

$$(3.12) \quad d\bar{Y}_t = \frac{\lambda(\bar{Y}_t) - \bar{Y}_t}{t+1} dt.$$

We assume that $\bar{Y}_0 = Y_0$. We have $Y_t^2 = \frac{N_t^2}{(t+1)^2}$. Applying Itô's formula for jump processes,

$$(3.13) \quad \begin{aligned} dY_t &= -\frac{Y_t}{t+1} dt + \frac{dN_t}{t+1}, \\ d(Y_t^2) &= -\frac{2N_t^2}{(t+1)^3} dt + \frac{(2N_{t-} + 1)dN_t}{(t+1)^2} = \frac{-2Y_t^2}{t+1} dt + \left(\frac{2Y_{t-}}{t+1} + \frac{1}{(t+1)^2} \right) dN_t, \\ d\bar{Y}_t &= \frac{\lambda(\bar{Y}_t) - \bar{Y}_t}{t+1} dt, \\ d(\bar{Y}_t)^2 &= 2\bar{Y}_t \frac{\lambda(\bar{Y}_t) - \bar{Y}_t}{t+1} dt. \end{aligned}$$

Therefore, we can compute that

$$(3.14) \quad \begin{aligned} d(Y_t - \bar{Y}_t)^2 = & -\frac{2(Y_t - \bar{Y}_t)^2}{t+1}dt + 2\frac{\bar{Y}_t}{t+1}[\lambda(\bar{Y}_t) - \lambda(Y_t)]dt - 2\frac{Y_t}{t+1}[\lambda(\bar{Y}_t) - \lambda(Y_t)]dt \\ & + \frac{\lambda(Y_t)}{(t+1)^2}dt + \left[\left(\frac{2Y_{t-}}{t+1} + \frac{1}{(t+1)^2} \right) - \frac{2\bar{Y}_t}{t+1} \right] dM_t. \end{aligned}$$

Define $m(t) := \mathbb{E}[(Y_t - \bar{Y}_t)^2]$. Hence,

$$(3.15) \quad \begin{aligned} \frac{dm(t)}{dt} = & -\frac{2m(t)}{t+1} + \frac{2}{t+1} \mathbb{E}[(\bar{Y}_t - Y_t)(\lambda(\bar{Y}_t) - \lambda(Y_t))] + \frac{\mathbb{E}[\lambda(Y_t)]}{(t+1)^2} \\ \leq & 2(\alpha - 1)\frac{m(t)}{t+1} + \frac{\mathbb{E}[\lambda(Y_t)]}{(t+1)^2}. \end{aligned}$$

Since $\lambda(\cdot)$ is α -Lipschitz for some $0 < \alpha < 1$, there exist some $c_1, c_2 > 0$ and $c_2 < 1$ so that $\lambda(z) \leq c_1 + c_2 z$. Under this condition, by Proposition 14, $\frac{\mathbb{E}[N_t]}{t+1} \leq K$ uniformly in t for some $K > 0$ and $\mathbb{E}[\lambda_t] = \mathbb{E}[\lambda(\frac{N_{t-}}{t+1})] \leq c_1 + c_2 K$ uniformly in t . Therefore,

$$(3.16) \quad \frac{dm(t)}{dt} \leq \frac{2(\alpha - 1)}{t+1}m(t) + \frac{c_1 + c_2 K}{(t+1)^2}, \quad m(0) = 0.$$

It is easy to verify that the solution to the ODE

$$(3.17) \quad \frac{dm(t)}{dt} = \frac{2(\alpha - 1)}{t+1}m(t) + \frac{c_1 + c_2 K}{(t+1)^2}, \quad m(0) = 0,$$

when $\alpha \neq \frac{1}{2}$ is given by

$$(3.18) \quad m(t) = \frac{c_1 + c_2 K}{1 - 2\alpha} \left[\frac{1}{t+1} - \frac{1}{(t+1)^{2(1-\alpha)}} \right].$$

When $\alpha = \frac{1}{2}$, the solution is given by

$$(3.19) \quad m(t) = \frac{(c_1 + c_2 K) \log(t+1)}{t+1}.$$

Therefore, we proved (2.5) and it is clear that $m(t) \rightarrow \infty$ as $t \rightarrow \infty$. Since \bar{Y}_t converges to the unique fixed point deterministically, we conclude that $\frac{N_t}{t+1}$ converges to the same value in the $L^2(\mathbb{P})$ norm. \square

Proof of Theorem 4. For any x that is not a fixed point of the equation $x = \lambda(x)$, then either $x > \lambda(x)$ or $x < \lambda(x)$. Let us assume without loss of generality that $\lambda(x) > x$. By continuity of $\lambda(x)$, there exists a sufficiently small $\epsilon > 0$ such that $\lambda(x - \epsilon) > x + \epsilon$. We claim that

$$(3.20) \quad \mathbb{P} \left(\omega : \exists N(\omega), \forall t \geq N(\omega), x - \epsilon < \frac{N_{t-}}{t+1} < x + \epsilon \right) = 0.$$

Notice that if $x - \epsilon < \frac{N_{t-}}{t+1} < x + \epsilon$ for any $t \geq N(\omega)$, then, from the monotonicity of the function $\lambda(x)$, we have

$$(3.21) \quad \lambda_t = \lambda \left(\frac{N_{t-}}{t+1} \right) \in [\lambda(x - \epsilon), \lambda(x + \epsilon)],$$

which is bounded below by $\lambda(x - \epsilon)$. But for a standard Poisson process N_t with constant intensity $\lambda(x - \epsilon)$, $\frac{N_t}{t+1} \rightarrow \lambda(x - \epsilon) > x + \epsilon$ almost surely, which implies (3.20). And (3.20) implies that

$$(3.22) \quad \lim_{t \rightarrow \infty} \mathbb{P} \left(\frac{N_t}{t} \in (x - \epsilon, x + \epsilon) \right) = \lim_{t \rightarrow \infty} \mathbb{P} \left(\frac{N_{t-}}{t+1} \in (x - \epsilon, x + \epsilon) \right) = 0.$$

Note that for the ϵ above, it depends on x . Now, consider any $I = [a, b]$ not containing any fixed point of $x = \lambda(x)$ and assume $\lambda(x) > x$ for any $x \in [a, b]$. Since $x \mapsto \lambda(x)$ is continuous, there exists $\epsilon > 0$ sufficiently small so that uniformly in $x \in [a, b]$, $\lambda(x - \epsilon) > x + \epsilon$. Hence the proof is complete. \square

Proof of Theorem 5. If x^* is the unique and stable fixed point of $\lambda(x)$, then $\frac{N_t}{t} \rightarrow x^*$ as $t \rightarrow \infty$ by using the previous result. Therefore, it is sufficient to show the following lemma.

Lemma 29. *Given that x^* a stable fixed point of $\lambda(x)$, there exists $\epsilon > 0$ and $t_0 > 0$ such that conditional on $\frac{N_{t_0}}{t_0} \in (x^* - \epsilon, x^* + \epsilon)$,*

$$(3.23) \quad \mathbb{P} \left(\frac{N_t}{t} \in (x^* - 2\epsilon, x^* + 2\epsilon), \forall t \geq t_0 \right) > 0$$

If Lemma 29 holds, as long as $\frac{N_t}{t} \in (x^* - 2\epsilon, x^* + 2\epsilon)$, $\forall t \geq t_0$, we can modify $\lambda(x)$ outside $(x^* - 2\epsilon, x^* + 2\epsilon)$ so that x^* is the unique fixed point. In addition, taking into account that $\frac{N_t}{t}$ is Markov and the event that $\frac{N_{t_0}}{t_0} \in (x^* - \epsilon, x^* + \epsilon)$ for some $t_0 > 0$ has positive probability, the proof is completed. \square

Proof of Lemma 29. Because x^* is a stable fixed point, $|\lambda'(x^*)| < 1$, and we can find $\epsilon > 0$ and $\delta > 0$, such that $x^* - \epsilon < \lambda(x) < x^* + \epsilon$ for $x \in (x^* - \epsilon - \delta, x^* + \epsilon + \delta)$.

Define a stopping time $\tau = \inf\{t \geq t_0 : \frac{N_t}{t} \notin (x^* - \epsilon - \delta, x^* + \epsilon + \delta)\}$. By using the coupling argument, we can construct two Poisson processes N_t^1 and N_t^2 with the intensity $\lambda_1 = x^* - \epsilon$ and $\lambda_2 = x^* + \epsilon$, respectively, such that

$$(3.24) \quad N_t^1 \leq N_t \leq N_t^2, \quad t \leq \tau$$

almost surely. Therefore, if $\tau_1 = \inf\{t \geq t_0 : \frac{N_t^1}{t} \geq x^* + \epsilon + \delta\}$ and $\tau_2 = \inf\{t \geq t_0 : \frac{N_t^2}{t} \leq x^* - \epsilon - \delta\}$, we have $\tau \geq \tau_1 \wedge \tau_2$ almost surely and

$$\mathbb{P}(t_0 \leq \tau < \infty) \leq \mathbb{P}(t_0 \leq \tau_1 \wedge \tau_2 < \infty) \leq \mathbb{P}(t_0 \leq \tau_1 < \infty) + \mathbb{P}(t_0 \leq \tau_2 < \infty)$$

By the strong law of the large numbers $\mathbb{P}(t_0 \leq \tau_1 < \infty), \mathbb{P}(t_0 \leq \tau_2 < \infty) \rightarrow 0$ as $t_0 \rightarrow \infty$. Finally, by letting $\delta < \epsilon$,

$$\begin{aligned} \mathbb{P} \left(\frac{N_t}{t} \in (x^* - 2\epsilon, x^* + 2\epsilon), \forall t \geq t_0 \right) &\geq 1 - \mathbb{P}(t_0 \leq \tau < \infty) \\ &\geq 1 - \mathbb{P}(t_0 \leq \tau_1 < \infty) - \mathbb{P}(t_0 \leq \tau_2 < \infty) > 0 \end{aligned}$$

for sufficiently large t_0 and we complete the proof. \square

Proof of Theorem 6. Let x^* be a strictly unstable fixed point. There exists a sufficiently small neighborhood C containing x^* so that for any $x \in C$ and $x \geq x^*$, we have $\lambda(x) \geq x$ and for any $x \in C$ and $x \leq x^*$, we have $\lambda(x) \leq x$. Let $Y_t = \frac{N_t}{t+1}$, where N_t is the simple point process with intensity $\lambda(\frac{N_{t-} + \gamma}{t+1})$ at time t and let $\tilde{Y}_t = \frac{\tilde{N}_t}{t+1}$, where \tilde{N}_t is the simple point process with intensity $\frac{\tilde{N}_{t-} + \gamma}{t+1}$ at time t .

Finally, we introduce the process $\hat{Y}_t = \frac{\hat{N}_t + \gamma}{t+1}$ so that the intensity of \hat{N}_t is $\lambda(\frac{\hat{N}_t + \gamma}{t+1})$ when $\hat{Y}_t \notin C$ and the intensity of \hat{N}_t is $\frac{\hat{N}_t + \gamma}{t+1}$ when $\hat{Y}_t \in C$. Since $\lambda(x) \geq x$ and for any $x \in C$ and $x \geq x^*$, and $\lambda(x) \leq x$ for any $x \in C$ and $x \leq x^*$, it is clear that $\mathbb{P}(\lim_{t \rightarrow \infty} Y_t = x^*) \leq \mathbb{P}(\lim_{t \rightarrow \infty} \hat{Y}_t = x^*)$. On the other hand, for the process \tilde{Y}_t , we proved in Proposition 21 that $\mathbb{P}(\lim_{t \rightarrow \infty} \tilde{Y}_t = \chi(\gamma)) = 1$, where $\chi(\gamma)$ is a gamma random variable with shape γ and scale 1. Therefore, for any $x \in \mathbb{R}^+$ and hence x^* , $\mathbb{P}(\lim_{t \rightarrow \infty} \tilde{Y}_t = x) = 0$. Since \hat{Y}_t shares the same dynamics as \tilde{Y}_t in C , we have $\mathbb{P}(\lim_{t \rightarrow \infty} \hat{Y}_t = x^*) = 0$, which implies that $\mathbb{P}(\lim_{t \rightarrow \infty} Y_t = x^*) = 0$. \square

3.2. Proofs of Results in Section 2.2.

Proof of Theorem 7. To prove the lower bound, it suffices to prove that (since we have the superexponential estimates (3.39) and (3.40))

$$(3.25) \quad \liminf_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P} \left(\frac{N_{\alpha_1 T}}{T} \in B_\epsilon(x_1), \dots, \frac{N_{\alpha_n T}}{T} \in B_\epsilon(x_n) \right) \geq -I(f),$$

where $B_\epsilon(x_i)$ are open balls centered at x_i with radius $\epsilon > 0$ and $f(\alpha)$, $0 \leq \alpha \leq 1$ is piecewise linear such that $f(\alpha_j) = x_j$ for any j , where $0 = \alpha_0 < \alpha_1 < \dots < \alpha_n = 1$.

We tilt λ_s to $\frac{x_j - x_{j-1}}{\alpha_j - \alpha_{j-1}}$ for $\alpha_{j-1}T < s \leq \alpha_j T$. Under the new measure, let us use induction. Assume that $\frac{N_{\alpha_{j-1}T}}{T} \rightarrow x_{j-1}$. Then, if we do not tilt on $[\alpha_{j-1}t, \alpha_j t]$, then, we get

$$(3.26) \quad \frac{N_{\alpha_j T}}{T} = \frac{N_{\alpha_{j-1}T}}{T} + \frac{N[\alpha_{j-1}T, \alpha_j T]}{T} \rightarrow x_{j-1} + \frac{x_j - x_{j-1}}{\alpha_j - \alpha_{j-1}}(\alpha_j - \alpha_{j-1}) = x_j.$$

Let $\hat{\mathbb{P}}$ denote the tilted probability measure and

$$(3.27) \quad A_T := \left\{ \frac{N_{\alpha_1 T}}{T} \in B_\epsilon(x_1), \dots, \frac{N_{\alpha_n T}}{T} \in B_\epsilon(x_n) \right\}.$$

The tilted probability measure $\hat{\mathbb{P}}$ is absolutely continuous w.r.t. \mathbb{P} and we have the following Girsanov formula, (For the theory of absolute continuity for point processes and its Girsanov formula, we refer to Lipster and Shiryaev [21].)

$$(3.28) \quad \left. \frac{d\hat{\mathbb{P}}}{d\mathbb{P}} \right|_{\mathcal{F}_T} = \exp \left\{ \sum_{j=1}^n \int_{\alpha_{j-1}T}^{\alpha_j T} \log \left(\frac{\frac{x_j - x_{j-1}}{\alpha_j - \alpha_{j-1}}}{\lambda_s} \right) dN_s + \int_{\alpha_{j-1}T}^{\alpha_j T} \lambda_s - \left(\frac{x_j - x_{j-1}}{\alpha_j - \alpha_{j-1}} \right) ds \right\}.$$

By Jensen's inequality, we have

$$(3.29) \quad \begin{aligned} \frac{1}{T} \log \mathbb{P}(A_T) &= \frac{1}{T} \log \int_{A_T} \frac{d\mathbb{P}}{d\hat{\mathbb{P}}} d\hat{\mathbb{P}} \\ &= \frac{1}{T} \log \hat{\mathbb{P}}(A_T) + \frac{1}{T} \log \left[\frac{1}{\hat{\mathbb{P}}(A_T)} \int_{A_T} \frac{d\mathbb{P}}{d\hat{\mathbb{P}}} d\hat{\mathbb{P}} \right] \\ &\geq \frac{1}{T} \log \hat{\mathbb{P}}(A_T) - \frac{1}{\hat{\mathbb{P}}(A_T)} \cdot \frac{1}{T} \hat{\mathbb{E}} \left[1_{A_T} \log \frac{d\hat{\mathbb{P}}}{d\mathbb{P}} \right]. \end{aligned}$$

Hence, we have

(3.30)

$$\begin{aligned}
& \liminf_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P} \left(\frac{N_{\alpha_1 T}}{T} \in B_\epsilon(x_1), \dots, \frac{N_{\alpha_n T}}{T} \in B_\epsilon(x_n) \right) \\
& \geq - \lim_{T \rightarrow \infty} \frac{1}{T} \hat{\mathbb{E}} \left[\sum_{j=1}^n \int_{\alpha_{j-1} T}^{\alpha_j T} \log \left(\frac{\frac{x_j - x_{j-1}}{\alpha_j - \alpha_{j-1}}}{\lambda_s} \right) dN_s + \int_{\alpha_{j-1} T}^{\alpha_j T} \lambda_s - \left(\frac{x_j - x_{j-1}}{\alpha_j - \alpha_{j-1}} \right) ds \right] \\
& = - \sum_{j=1}^n \log \left(\frac{\frac{x_j - x_{j-1}}{\alpha_j - \alpha_{j-1}}}{\lambda(x_j/\alpha_j)} \right) (x_j - x_{j-1}) + (\alpha_j - \alpha_{j-1}) \lambda(x_j/\alpha_j) - (x_j - x_{j-1}) \\
& = -I(f).
\end{aligned}$$

To prove the upper bound for compact sets, it is sufficient to prove that for any piecewise linear $f \in \mathcal{AC}_0[0, 1]$,

$$(3.31) \quad \limsup_{\epsilon \rightarrow 0} \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{P} \left(\frac{N \cdot t}{t} \in B_\epsilon(f) \right) \leq -I(f).$$

To prove the upper bound for closed sets instead of compact sets, one needs to prove some superexponential estimates which will be discussed later.

Notice that

$$(3.32) \quad 1 = \mathbb{E} \left[e^{\int_0^1 \Phi(\alpha) dN_{\alpha T} - \int_0^1 (e^{\Phi(\alpha)} - 1) \lambda \left(\frac{N_{\alpha T}}{1 + \alpha T} \right) d(\alpha T)} \right],$$

for any bounded function Φ . That is because for any $f(s, \omega)$ which is bounded, progressively measurable and \mathcal{F}_t -predictable,

$$(3.33) \quad \exp \left\{ \int_0^t f(s, \omega) dN_s - \int_0^t (e^{f(s, \omega)} - 1) \lambda_s ds \right\},$$

is a martingale.

Let us choose the test functions Φ as a step function and assume that there exists a sequence $0 = \alpha_0 < \alpha_1 < \dots < \alpha_M = 1$ such that $\Phi(\alpha) = \beta_j$ for any $\alpha_{j-1} < \alpha < \alpha_j$, $1 \leq j \leq M$.

For $\frac{N \cdot T}{T} \in B_\epsilon(f)$, we have

$$\begin{aligned}
(3.34) \quad & \left| \int_{\alpha_{j-1}}^{\alpha_j} \Phi(\alpha) d \left(\frac{dN_{\alpha T}}{T} \right) - \int_{\alpha_{j-1}}^{\alpha_j} \Phi(\alpha) f'(\alpha) d\alpha \right| \\
& = \left| \beta_j \frac{N_{\alpha_j T} - N_{\alpha_{j-1} T}}{T} - \beta_j (f(\alpha_j) - f(\alpha_{j-1})) \right| \leq 2|\beta_j| \epsilon.
\end{aligned}$$

Moreover,

$$(3.35) \quad \left| \int_0^1 (e^\Phi - 1) \lambda \left(\frac{N_{\alpha T}}{1 + \alpha T} \right) d\alpha - \int_0^1 (e^\Phi - 1) \lambda \left(\frac{f(\alpha)}{\alpha} \right) d\alpha \right| \leq \sup_{0 \leq \alpha \leq 1} |e^\Phi - 1| \gamma \cdot \epsilon.$$

Hence, by Chebychev's inequality, we have

$$\begin{aligned}
 (3.36) \quad & \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P} \left(\frac{N_{\alpha T}}{T} \in B_\epsilon(f), 0 \leq \alpha \leq 1 \right) \\
 & \leq - \left\{ \int_0^1 \Phi(\alpha) f'(\alpha) d\alpha - \int_0^1 (e^\Phi - 1) \lambda \left(\frac{f(\alpha)}{\alpha} \right) d\alpha \right\} \\
 & \quad + 2\epsilon \sum_{j=1}^M |\beta_j| + \sup_{0 \leq \alpha \leq 1} |e^\Phi - 1| \gamma \cdot \epsilon.
 \end{aligned}$$

Hence, we have

$$\begin{aligned}
 (3.37) \quad & \limsup_{\epsilon \rightarrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P} \left(\frac{N_{\epsilon T}}{T} \in B_\epsilon(f) \right) \\
 & \leq - \left\{ \int_0^1 \Phi(\alpha) f'(\alpha) d\alpha - \int_0^1 (e^\Phi - 1) \lambda \left(\frac{f(\alpha)}{\alpha} \right) d\alpha \right\}.
 \end{aligned}$$

We can optimize over Φ by choosing $\Phi = \log(\frac{f'(\alpha)}{\lambda(f(\alpha)/\alpha)})$. Hence, we proved that

$$(3.38) \quad \limsup_{\epsilon \rightarrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P} \left(\frac{N_{\epsilon T}}{T} \in B_\epsilon(f) \right) \leq -I(f).$$

Finally, we need to obtain the superexponential estimates in order to prove the upper bound for closed sets instead of compact sets in the topology of uniform convergence. This is not difficult because the jump rate $\lambda(\cdot) \leq C_0$. We have the following superexponential estimates,

$$\begin{aligned}
 (3.39) \quad & \limsup_{K \rightarrow \infty} \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P} \left(\sup_{\alpha \in [0,1]} \frac{N_{\alpha T}}{T} \geq K \right) \\
 & = \limsup_{K \rightarrow \infty} \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P}(N_T \geq KT) = -\infty,
 \end{aligned}$$

and for any $\delta > 0$

$$\begin{aligned}
 (3.40) \quad & \limsup_{\epsilon \rightarrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P} \left(\sup_{|\alpha - \beta| \leq \epsilon, 0 \leq \alpha, \beta \leq 1} \left| \frac{N_{\alpha T}}{T} - \frac{N_{\beta T}}{T} \right| \geq \delta \right) \\
 & \leq \limsup_{\epsilon \rightarrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P} \left(\exists 1 \leq j \leq [1/\epsilon] : N[j\epsilon T, (j+1)\epsilon T] \geq \frac{T\delta}{2} \right) \\
 & \leq \limsup_{\epsilon \rightarrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log [1/\epsilon] \mathbb{P} \left(N^+[0, \epsilon T] > \frac{T\delta}{2} \right),
 \end{aligned}$$

where N^+ is the Poisson process with constant rate C_0 . Applying Chebychev's inequality and setting $\theta = \log(\frac{1+\epsilon}{\epsilon})$, we have

$$\begin{aligned}
 (3.41) \quad & \limsup_{\epsilon \rightarrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log [1/\epsilon] \mathbb{P} \left(N^+[0, \epsilon T] > \frac{T\delta}{2} \right) \\
 & \leq \limsup_{\epsilon \rightarrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log [1/\epsilon] e^{C_0(\epsilon^\theta - 1)\epsilon T - \theta\delta T/2} \\
 & \leq \limsup_{\epsilon \rightarrow 0} \left\{ C_0 - \log \left(\frac{1+\epsilon}{\epsilon} \right) \frac{\delta}{2} \right\} = -\infty.
 \end{aligned}$$

Hence, we have, for any closet set C ,

$$(3.42) \quad \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P} \left(\frac{N_{\cdot T}}{T} \in C \right) \leq - \inf_{f \in C} I(f).$$

□

Proof of Proposition 12. We first note that $I(x)$ is a good rate function and therefore there must be a function $f \in \mathcal{AC}_0^+[0, 1]$ with $f(0) = 0$ and $f(1) = x$ such that

$$(3.43) \quad I(x) = \int_0^1 L(\alpha, f(\alpha), f'(\alpha)) d\alpha = 0.$$

We know that $L(\alpha, f(\alpha), f'(\alpha)) \geq 0$ and $L(\alpha, f(\alpha), f'(\alpha)) = 0$ if and only if

$$(3.44) \quad f'(\alpha) = \lambda \left(\frac{f(\alpha)}{\alpha} \right).$$

As the limit $\alpha \rightarrow 0^+$, we have

$$f'(0) = \lim_{\alpha \rightarrow 0^+} \lambda \left(\frac{f(\alpha)}{\alpha} \right) = \lim_{\alpha \rightarrow 0^+} \lambda \left(\frac{f(\alpha) - f(0)}{\alpha} \right) = \lambda(f'(0)).$$

Thus, $f'(0)$ must be a fixed point x^* of $\lambda(x)$, i.e., $x^* = \lambda(x^*)$. Then we discretize (3.44) by using the Euler method:

$$(3.45) \quad \begin{cases} f_1 = f_0 + \Delta\alpha \lambda(f'(0)), & f_0 = 0 \\ f_{n+1} = f_n + \Delta\alpha \lambda \left(\frac{f_n}{n\Delta\alpha} \right), & 1 \leq n < N = \frac{1}{\Delta\alpha}. \end{cases}$$

By using the fact that $f'(0) = x^* = \lambda(x^*)$, it is easy to see that $f_n = x^* n \Delta\alpha$ for all n and $f_N = x^* N \Delta\alpha = x^*$. When $\Delta\alpha \rightarrow 0$, $\{f_n\}_{n=0}^N$ obtained by (3.45) converges to the solution of (3.44). Therefore, as $\Delta\alpha \rightarrow 0$, $x^* = f_N \rightarrow f(1) = x$ so $x = x^*$ which is a fixed point of $\lambda(x)$. In addition, $f(\alpha)$ must be linear: $f(\alpha) = \alpha x$, for $\alpha \in [0, 1]$. □

3.3. Proofs of Results in Section 2.3.

Proof of Proposition 14. (i) Let us recall that

$$(3.46) \quad \begin{aligned} dY_t &= -\frac{Y_t}{t+1} dt + \frac{dN_t}{t+1}, \\ d(Y_t^2) &= \frac{-2Y_t^2}{t+1} dt + \left(\frac{2Y_{t-}}{t+1} + \frac{1}{(t+1)^2} \right) dN_t. \end{aligned}$$

Let us assume that $\lambda(z) = \alpha z + \beta$, where $\alpha, \beta > 0$ and $\alpha < 1$ so that there is a unique fixed point to the equation $z = \lambda(z)$ at $z^* = \frac{\beta}{1-\alpha}$.

Let $m_1(t) = \mathbb{E}[Y_t]$ and assume $Y_0 = 0$, then,

$$(3.47) \quad \frac{dm_1(t)}{dt} = \frac{(\alpha - 1)m_1(t) + \beta}{t+1}, \quad m_1(0) = 0,$$

which implies that

$$(3.48) \quad m_1(t) = \frac{\beta}{1-\alpha} [1 - (t+1)^{\alpha-1}],$$

which yields (2.13).

(ii) Let $m_2(t) = \mathbb{E}[Y_t^2]$. Then,

$$(3.49) \quad \frac{dm_2(t)}{dt} = \frac{2(\alpha-1)m_2(t)}{t+1} + \frac{2\beta m_1(t)}{t+1} + \frac{\beta + \alpha m_1(t)}{(t+1)^2}, \quad m_2(0) = 0.$$

Therefore,

$$(3.50) \quad \begin{aligned} \frac{dm_2(t)}{dt} &= \frac{2(\alpha-1)m_2(t)}{t+1} + \frac{2\beta^2}{1-\alpha} \left[\frac{1}{t+1} - \frac{1}{(t+1)^{2-\alpha}} \right] \\ &\quad + \frac{\frac{\beta}{1-\alpha}}{(t+1)^2} - \frac{\beta\alpha}{1-\alpha} \frac{1}{(t+1)^{3-\alpha}}, \quad m_2(0) = 0. \end{aligned}$$

Consider $m_2(t) = \frac{C_1}{t+1} + \frac{C_2}{(t+1)^{2-\alpha}} + C_3 + \frac{C_4}{(t+1)^{1-\alpha}} + \frac{C_5}{(t+1)^{2(1-\alpha)}}$. Then,

$$(3.51) \quad \begin{aligned} & -\frac{C_1}{(t+1)^2} - (2-\alpha)\frac{C_2}{(t+1)^{3-\alpha}} - (1-\alpha)\frac{C_4}{(t+1)^{2-\alpha}} + 2(\alpha-1)\frac{C_5}{(t+1)^{3-2\alpha}} \\ &= \frac{2(\alpha-1)C_1}{(t+1)^2} + \frac{2(\alpha-1)C_2}{(t+1)^{3-\alpha}} + \frac{2(\alpha-1)C_3}{t+1} \\ &\quad + \frac{2\beta^2}{1-\alpha} \frac{1}{t+1} + \frac{\frac{\beta}{1-\alpha}}{(t+1)^2} - \frac{\beta\alpha}{1-\alpha} \frac{1}{(t+1)^{3-\alpha}} \\ &\quad + 2(\alpha-1)\frac{C_4}{(t+1)^{2-\alpha}} - \frac{2\beta^2}{1-\alpha} \frac{1}{(t+1)^{2-\alpha}} + 2(\alpha-1)\frac{C_5}{(t+1)^{3-2\alpha}}. \end{aligned}$$

Therefore, we have

$$(3.52) \quad \begin{aligned} C_1 &= \frac{\beta}{(1-2\alpha)(1-\alpha)}, \\ C_2 &= \frac{\beta}{1-\alpha}, \\ C_3 &= \frac{\beta^2}{(1-\alpha)^2}, \\ C_4 &= \frac{-2\beta^2}{(1-\alpha)^2}. \end{aligned}$$

Finally, since $m_2(0) = C_1 + C_2 + C_3 + C_4 + C_5 = 0$, we have

$$(3.53) \quad C_5 = -C_1 - C_2 - C_3 - C_4 = -\frac{\beta}{1-\alpha} \left[1 - \frac{\beta}{1-\alpha} + \frac{1}{1-2\alpha} \right].$$

Therefore,

(3.54)

$$\begin{aligned}
\text{Var}[N_t] &= (t+1)^2 [m_2(t) - m_1(t)^2] \\
&= (t+1)^2 \left[\frac{\frac{\beta}{(1-2\alpha)(1-\alpha)}}{t+1} + \frac{\frac{\beta}{1-\alpha}}{(t+1)^{2-\alpha}} + \frac{\beta^2}{(1-\alpha)^2} \right. \\
&\quad + \frac{\frac{-2\beta^2}{(1-\alpha)^2}}{(t+1)^{1-\alpha}} + \frac{-\frac{\beta}{1-\alpha} \left[1 - \frac{\beta}{1-\alpha} + \frac{1}{1-2\alpha} \right]}{(t+1)^{2(1-\alpha)}} \\
&\quad \left. - \frac{\beta^2}{(1-\alpha)^2} - \frac{\beta^2}{(1-\alpha)^2} \frac{1}{(t+1)^{2(1-\alpha)}} + 2 \frac{\beta^2}{(1-\alpha)^2} \frac{1}{(t+1)^{1-\alpha}} \right] \\
&= (t+1)^2 \left[\frac{\frac{\beta}{(1-2\alpha)(1-\alpha)}}{t+1} + \frac{\frac{\beta}{1-\alpha}}{(t+1)^{2-\alpha}} + \frac{-\frac{2\beta}{1-2\alpha}}{(t+1)^{2(1-\alpha)}} \right].
\end{aligned}$$

Hence, we proved (2.16).

For $0 < \alpha < \frac{1}{2}$, from (2.16), it is easy to check that

$$(3.55) \quad \lim_{t \rightarrow \infty} \frac{\text{Var}[N_t]}{t} = \frac{\beta}{(1-2\alpha)(1-\alpha)}.$$

For $\alpha > \frac{1}{2}$, from (2.16), it is easy to check that

$$(3.56) \quad \lim_{t \rightarrow \infty} \frac{\text{Var}[N_t]}{t^{2\alpha}} = \frac{2\beta}{2\alpha-1}.$$

□

Proof of Theorem 15. Cox and Grimmett [8] has a central limit theorem for associated random variables. For a sequence of associated random variables $(X_n)_{n=1}^\infty$, if it satisfies

- (i) $\text{Var}[X_n] \geq c_1$ and $\mathbb{E}[|X_n|^3] \leq c_2$.
- (ii) $\sum_{j: |n-j| \geq r} \text{Cov}(X_j, X_n) \leq u(r) \rightarrow 0$ as $r \rightarrow \infty$.

Then, $\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\text{Var}[S_n]}} \rightarrow N(0, 1)$ in distribution as $n \rightarrow \infty$.

For self-exciting point processes, a new jump will increase the intensity which will help generate more jumps. Under the assumption $\lambda(\cdot)$ is increasing, our model is in the class of self-exciting point processes studied by Kwieciński and Szekli [20] and $(N(n, n+1))_{n=0}^\infty$ are associated random variables.

One can use the formulas in Proposition 14 to show (i) directly. Alternatively, we can observe that for any $T > 0$, the compensator of $N[t, t+T]$ is $\int_t^{t+T} \lambda_s ds$ which converges to $\frac{\beta}{1-\alpha} T$ as $t \rightarrow \infty$. Hence $N[t, t+T]$, $T > 0$ converges to a standard Poisson process \bar{N} with parameter $\frac{\beta}{1-\alpha}$ as $t \rightarrow \infty$. Thus, $\text{Var}[N[t, t+1]] \rightarrow \text{Var}[\bar{N}[0, 1]]$, $\mathbb{E}[N[t, t+1]^3] \rightarrow \mathbb{E}[\bar{N}[0, 1]^3]$ as $t \rightarrow \infty$.

Note that by Proposition 19, for any $t > s + 1$,

$$\begin{aligned}
(3.57) \quad & \text{Cov}(N[t, t+1], N[s, s+1]) \\
&= \text{Cov}(N_{t+1}, N_{s+1}) - \text{Cov}(N_{t+1}, N_s) - \text{Cov}(N_t, N_{s+1}) + \text{Cov}(N_t, N_s) \\
&= (t+2)^\alpha \left[\frac{\beta}{(1-2\alpha)(1-\alpha)} (s+2)^{1-\alpha} + \frac{\beta}{1-\alpha} - \frac{2\beta}{1-2\alpha} (s+2)^\alpha \right] \\
&\quad - (t+2)^\alpha \left[\frac{\beta}{(1-2\alpha)(1-\alpha)} (s+1)^{1-\alpha} + \frac{\beta}{1-\alpha} - \frac{2\beta}{1-2\alpha} (s+1)^\alpha \right] \\
&\quad - (t+1)^\alpha \left[\frac{\beta}{(1-2\alpha)(1-\alpha)} (s+2)^{1-\alpha} + \frac{\beta}{1-\alpha} - \frac{2\beta}{1-2\alpha} (s+2)^\alpha \right] \\
&\quad + (t+1)^\alpha \left[\frac{\beta}{(1-2\alpha)(1-\alpha)} (s+1)^{1-\alpha} + \frac{\beta}{1-\alpha} - \frac{2\beta}{1-2\alpha} (s+1)^\alpha \right] \\
&= \frac{\beta}{(1-2\alpha)(1-\alpha)} [(t+2)^\alpha - (t+1)^\alpha] [(s+2)^{1-\alpha} - (s+1)^{1-\alpha}] \\
&\quad - \frac{2\beta}{1-2\alpha} [(t+2)^\alpha - (t+1)^\alpha] [(s+2)^\alpha - (s+1)^\alpha] \\
&\leq \frac{\beta}{(1-2\alpha)(1-\alpha)} [(t+2)^\alpha - (t+1)^\alpha] [(s+2)^{1-\alpha} - (s+1)^{1-\alpha}] \\
&\leq \frac{\beta}{(1-2\alpha)(1-\alpha)} \frac{\alpha}{(t+1)^{1-\alpha}} \frac{1-\alpha}{(s+1)^\alpha} \\
&\leq \frac{\alpha\beta}{1-2\alpha} \frac{1}{(t+1)^{1-\alpha}}.
\end{aligned}$$

This proved (ii). Since $\mathbb{E}[N[t, t+1]]$ is uniformly bounded in t , discrete time CLT can be replaced by continuous time CLT and we have $\frac{N_t - \mathbb{E}[N_t]}{\sqrt{\text{Var}[N_t]}} \rightarrow N(0, 1)$ in distribution as $t \rightarrow \infty$. Finally, by the expressions of $\mathbb{E}[N_t]$ and $\text{Var}[N_t]$ in Proposition 14, we proved (2.19). \square

Proof of Proposition 16. Following the proof of Proposition 14, for $\alpha = \frac{1}{2}$, $m_2(t) = \mathbb{E}[Y_t^2]$,

$$\begin{aligned}
(3.58) \quad & \frac{dm_2(t)}{dt} = \frac{2(\alpha-1)m_2(t)}{t+1} + \frac{2\beta^2}{1-\alpha} \left[\frac{1}{t+1} - \frac{1}{(t+1)^{2-\alpha}} \right] \\
& \quad + \frac{\frac{\beta}{1-\alpha}}{(t+1)^2} - \frac{\beta\alpha}{1-\alpha} \frac{1}{(t+1)^{3-\alpha}}, \quad m_2(0) = 0.
\end{aligned}$$

Consider $m_2(t) = \frac{C_1}{t+1} + \frac{C_2}{(t+1)^{2-\alpha}} + C_3 + \frac{C_4}{(t+1)^{1-\alpha}} + \frac{C_5 \log(t+1)}{t+1}$ and use the initial condition $m_2(0) = 0$, we get

$$(3.59) \quad \begin{aligned} C_1 &= 4\beta^2 - 2\beta, \\ C_2 &= \frac{\beta}{1-\alpha} = 2\beta, \\ C_3 &= \frac{\beta^2}{(1-\alpha)^2} = 4\beta^2, \\ C_4 &= \frac{-2\beta^2}{(1-\alpha)^2} = -8\beta^2, \\ C_5 &= 2\beta. \end{aligned}$$

Therefore,

$$(3.60) \quad \begin{aligned} \text{Var}[N_t] &= (t+1)^2 [m_2(t) - (m_1(t))^2] \\ &= (t+1)^2 \left[\frac{4\beta^2 - 2\beta + 2\beta \log(t+1)}{t+1} + \frac{2\beta}{(t+1)^{3/2}} + 4\beta^2 + \frac{-8\beta^2}{(t+1)^{1/2}} \right. \\ &\quad \left. - 4\beta^2 - \frac{4\beta^2}{t+1} + \frac{8\beta^2}{(t+1)^{1/2}} \right] \\ &= (t+1)^2 \left[\frac{2\beta[\log(t+1) - 1]}{t+1} + \frac{2\beta}{(t+1)^{3/2}} \right]. \end{aligned}$$

□

Proof of Proposition 17. Let $m_1(t) = \mathbb{E}[Y_t]$ and assume $Y_0 = 0$, then,

$$(3.61) \quad \frac{dm_1(t)}{dt} = \frac{\beta}{t+1}, \quad m_1(0) = 0,$$

which yields that $m_1(t) = \beta \log(t+1)$.

Let $m_2(t) = \mathbb{E}[Y_t^2]$. Then, $m_2(0) = 0$,

$$(3.62) \quad \begin{aligned} \frac{dm_2(t)}{dt} &= \frac{2\beta m_1(t)}{t+1} + \frac{\beta + m_1(t)}{(t+1)^2} \\ &= \frac{2\beta^2 \log(t+1)}{t+1} + \frac{\beta + \beta \log(t+1)}{(t+1)^2}, \end{aligned}$$

which implies that

$$(3.63) \quad m_2(t) = \beta^2 [\log(t+1)]^2 - \frac{\beta \log(t+1)}{t+1} + 2\beta \left[1 - \frac{1}{t+1} \right].$$

Therefore,

$$(3.64) \quad \begin{aligned} \text{Var}[N_t] &= (t+1)^2 [m_2(t) - (m_1(t))^2] \\ &= (t+1)^2 \left[-\frac{\beta \log(t+1)}{t+1} + 2\beta \left[1 - \frac{1}{t+1} \right] \right] \\ &= -\beta(t+1) \log(t+1) + 2\beta t(t+1). \end{aligned}$$

□

Proof of Corollary 18. The proof follows from the results in Proposition 14, Proposition 16, Proposition 17 and Chebychev's inequality. □

Proof of Proposition 19. For any $t > s$,

$$(3.65) \quad \mathbb{E}[N_t N_s] = \mathbb{E}[N_s^2] + \mathbb{E} \left[N_s \int_s^t \lambda_u du \right] = \mathbb{E}[N_s^2] + \int_s^t \left(\beta + \alpha \frac{\mathbb{E}[N_u N_s]}{u+1} \right) du.$$

Let $m(t, s) := \mathbb{E}[N_t N_s]$. Then

$$(3.66) \quad \frac{\partial m}{\partial t} = \beta \mathbb{E}[N_s] + \frac{\alpha}{t+1} m(t, s), \quad m(s, s) = \mathbb{E}[N_s^2],$$

which yields the solution when $\alpha \neq 1$,

$$(3.67) \quad m(t, s) = (t+1) \frac{\beta \mathbb{E}[N_s]}{1-\alpha} + (t+1)^\alpha \frac{\mathbb{E}[N_s^2] - \frac{\beta \mathbb{E}[N_s]}{1-\alpha} (s+1)}{(s+1)^\alpha}.$$

From the proofs of Proposition 14, for $\alpha \notin \{\frac{1}{2}, 1\}$,

$$(3.68) \quad \begin{aligned} \mathbb{E}[N_t] &= \frac{\beta}{1-\alpha} [(t+1) - (t+1)^\alpha], \\ \mathbb{E}[N_t^2] &= \frac{\beta}{(1-2\alpha)(1-\alpha)} (t+1) + \frac{\beta}{1-\alpha} (t+1)^\alpha + \frac{\beta^2}{(1-\alpha)^2} (t+1)^2 \\ &\quad - \frac{2\beta^2}{(1-\alpha)^2} (t+1)^{\alpha+1} - \frac{\beta}{1-\alpha} \left[1 - \frac{\beta}{1-\alpha} + \frac{1}{1-2\alpha} \right] (t+1)^{2\alpha}. \end{aligned}$$

Substituting them into (3.67) and using the identity $\text{Cov}[N_t, N_s] = \mathbb{E}[N_t N_s] - \mathbb{E}[N_t] \mathbb{E}[N_s]$, we get

$$(3.69) \quad \begin{aligned} \text{Cov}[N_t, N_s] &= \frac{\beta^2}{(1-\alpha)^2} (t+1) [(s+1) - (s+1)^\alpha] - \frac{\beta^2}{(1-\alpha)^2} [(s+1)^{2-\alpha} - (s+1)] (t+1)^\alpha \\ &\quad + (t+1)^\alpha \left[\frac{\beta}{(1-2\alpha)(1-\alpha)} (s+1)^{1-\alpha} + \frac{\beta}{1-\alpha} + \frac{\beta^2}{(1-\alpha)^2} (s+1)^{2-\alpha} \right. \\ &\quad \left. - \frac{2\beta^2}{(1-\alpha)^2} (s+1) - \frac{\beta}{1-\alpha} \left[1 - \frac{\beta}{1-\alpha} + \frac{1}{1-2\alpha} \right] (s+1)^\alpha \right] \\ &\quad - \frac{\beta^2}{(1-\alpha)^2} [(t+1) - (t+1)^\alpha] [(s+1) - (s+1)^\alpha] \\ &= (t+1)^\alpha \left[\frac{\beta}{(1-2\alpha)(1-\alpha)} (s+1)^{1-\alpha} + \frac{\beta}{1-\alpha} - \frac{2\beta}{1-2\alpha} (s+1)^\alpha \right]. \end{aligned}$$

For $\alpha = \frac{1}{2}$, from Proposition 16,

$$\begin{aligned} \mathbb{E}[N_t] &= 2\beta [(t+1) - (t+1)^{\frac{1}{2}}], \\ \mathbb{E}[N_t^2] &= (4\beta^2 - 2\beta)(t+1) + 2\beta(t+1)^{\frac{1}{2}} + 4\beta^2(t+1)^2 \\ &\quad - 8\beta^2(t+1)^{\frac{3}{2}} + 2\beta(t+1) \log(t+1). \end{aligned}$$

Hence, substituting these into (3.67), we get

(3.70)

$$\begin{aligned} \text{Cov}[N_t, N_s] &= m(t, s) - \mathbb{E}[N_t]\mathbb{E}[N_s] \\ &= (t+1)2\beta\mathbb{E}[N_s] + (t+1)^{\frac{1}{2}} \frac{\mathbb{E}[N_s^2] - 2\beta\mathbb{E}[N_s](s+1)}{(s+1)^{\frac{1}{2}}} \\ &\quad - 4\beta^2[(t+1) - (t+1)^{\frac{1}{2}}][(s+1) - (s+1)^{\frac{1}{2}}] \\ &= 2\beta[-(t+1)^{\frac{1}{2}}(s+1)^{\frac{1}{2}} + (t+1)^{\frac{1}{2}} + (t+1)^{\frac{1}{2}}(s+1)^{\frac{1}{2}}\log(s+1)]. \end{aligned}$$

□

Proof of Proposition 20. Let $m(t, s) = \mathbb{E}[N_t N_s]$. Following the proofs in Proposition 19,

$$(3.71) \quad \frac{\partial m}{\partial t} = \beta\mathbb{E}[N_s] + \frac{1}{t+1}m(t, s), \quad m(s, s) = \mathbb{E}[N_s^2],$$

which yields the solution

$$(3.72) \quad m(t, s) = \beta\mathbb{E}[N_s](t+1)\log(t+1) + (t+1)\frac{\mathbb{E}[N_s^2] - \beta\mathbb{E}[N_s](s+1)\log(s+1)}{s+1}.$$

From Proposition 17,

$$(3.73) \quad \begin{aligned} \mathbb{E}[N_t] &= \beta(t+1)\log(t+1), \\ \mathbb{E}[N_t^2] &= -\beta(t+1)\log(t+1) + 2\beta t(t+1) + \beta^2[(t+1)\log(t+1)]^2. \end{aligned}$$

Substituting this into (3.72) and using the identity $\text{Cov}[N_t, N_s] = \mathbb{E}[N_t N_s] - \mathbb{E}[N_t]\mathbb{E}[N_s]$, we get

$$\begin{aligned} (3.74) \quad \text{Cov}[N_t, N_s] &= \beta\mathbb{E}[N_s](t+1)\log(t+1) + (t+1)\frac{\mathbb{E}[N_s^2] - \beta\mathbb{E}[N_s](s+1)\log(s+1)}{s+1} \\ &\quad - \beta^2[(t+1)\log(t+1)][(s+1)\log(s+1)] \\ &= \frac{t+1}{s+1} [-\beta(s+1)\log(s+1) + 2\beta s(s+1) + \beta^2[(s+1)\log(s+1)]^2] \\ &\quad - \beta(t+1)\log(s+1)\beta(s+1)\log(s+1) \\ &= -\beta(t+1)\log(s+1) + 2\beta s(t+1). \end{aligned}$$

□

Proof of Proposition 21. Since

$$(3.75) \quad \mathbb{E}[N_t] = \mathbb{E} \left[\int_0^t \lambda_s ds \right] = \int_0^t \frac{\alpha(\mathbb{E}[N_s] + \gamma)}{s+1} ds.$$

By letting $g(t) := \mathbb{E}[N_t]$, it satisfies the ODE

$$(3.76) \quad g'(t) = \frac{\alpha(g(t) + \gamma)}{t+1}, \quad g(0) = 0,$$

which yields the solution $g(t) = \gamma[(t+1)^\alpha - 1]$. This is consistent with (2.36) and the variance of a negative binomial distribution. Next, let $h(t) = \mathbb{E}[N_t^2]$. Then

$$(3.77) \quad d(N_t^2) = (2N_{t-} + 1)dN_t,$$

and hence after taking expectations, $h(0) = 0$ and

$$(3.78) \quad \begin{aligned} h'(t) &= 2\alpha \frac{h(t)}{t+1} + \frac{2\gamma\alpha + \alpha}{t+1} g(t) + \frac{\gamma\alpha}{t+1} \\ &= \frac{2\alpha h(t)}{t+1} + \frac{2\gamma^2\alpha^2}{(t+1)^{1-\alpha}} + \gamma\alpha - \frac{2\gamma^2\alpha^2}{t+1}, \end{aligned}$$

which yields the solution

$$(3.79) \quad h(t) = \frac{\gamma^2\alpha^2}{\alpha^2} - \left[\frac{2\gamma^2\alpha^2}{\alpha^2} + \frac{\gamma\alpha}{\alpha} \right] (t+1)^\alpha + \left[\frac{\gamma^2\alpha^2}{\alpha^2} + \frac{\gamma\alpha}{\alpha} \right] (t+1)^{2\alpha}.$$

Hence

$$(3.80) \quad \text{Var}[N_t] = \gamma[(t+1)^{2\alpha} - (t+1)^\alpha].$$

This is consistent with (2.36) and the variance of a negative binomial distribution. Furthermore, by (2.37) and the properties of negative binomial distributions

(3.81)

$$\begin{aligned} \text{Cov}[N_t, N_s] &= \mathbb{E}[\mathbb{E}[N_t|N_s]N_s] - \mathbb{E}[N_t]\mathbb{E}[N_s] \\ &= \mathbb{E} \left[(N_s + \gamma) \left[\left(\frac{t+1}{s+1} \right)^\alpha - 1 \right] N_s + N_s^2 \right] - \mathbb{E}[N_t]\mathbb{E}[N_s] \\ &= \left[\left(\frac{t+1}{s+1} \right)^\alpha - 1 \right] [\gamma[(s+1)^{2\alpha} - (s+1)^\alpha] + \gamma^2[(s+1)^\alpha - 1]^2 + \gamma[(s+1)^\alpha - 1]] \\ &\quad + \gamma[(s+1)^{2\alpha} - (s+1)^\alpha] + \gamma^2[(s+1)^\alpha - 1]^2 - \gamma^2[(t+1)^\alpha - 1][(s+1)^\alpha - 1] \\ &= [(s+1)^\alpha - 1] \left[\gamma(t+1)^\alpha + (\gamma - \gamma^2) \left[\frac{(t+1)^\alpha}{(s+1)^\alpha} - 1 \right] \right]. \end{aligned}$$

Moreover,

$$(3.82) \quad d \left(\frac{N_t + \gamma}{(t+1)^\alpha} \right) = -\alpha \frac{N_t + \gamma}{(t+1)^{\alpha+1}} dt + \frac{dN_t}{(t+1)^\alpha} = \frac{dM_t}{(t+1)^\alpha}.$$

Hence, $\frac{N_t + \gamma}{(t+1)^\alpha}$ is a martingale and from (2.29) we have that $\sup_{t>0} \mathbb{E} \left[\frac{N_t + \gamma}{(t+1)^\alpha} \right] < \infty$.

Therefore, by martingale convergence theorem, $\frac{N_t}{t^\alpha} \rightarrow \chi(\alpha, \gamma)$, a.s. and in $L^2(\mathbb{P})$, for some random variable $\chi(\alpha, \gamma)$ which is finite a.s. and in $L^2(\mathbb{P})$ and possibly depends on parameters α and γ . Finally, by (2.36) and the formula for the Laplace transform of negative binomial distribution, for any $\theta > 0$,

$$(3.83) \quad \mathbb{E} \left[e^{-\theta \frac{N_t}{t^\alpha}} \right] = \left(\frac{\frac{1}{(t+1)^\alpha}}{1 - \left(1 - \frac{1}{(t+1)^\alpha} \right) e^{-\frac{\theta}{t^\alpha}}} \right)^\gamma \rightarrow \left(\frac{1}{1 + \theta} \right)^\gamma,$$

as $t \rightarrow \infty$. Hence $\chi(\alpha, \gamma)$ is independent of α and follows a gamma distribution with shape γ and scale 1. \square

Proof of Proposition 22. For any $T > 0$, $\lambda_t \geq \lambda(\frac{N_t + \gamma}{T+1})$ on $[0, T]$. Comparing with the pure-birth process, see e.g. Feller [12], it becomes clear that $\mathbb{P}(N(0, T] = \infty) > 0$. Moreover, to see $\mathbb{P}(\tau < \infty) < 1$, it suffices to notice that

$$(3.84) \quad \mathbb{P}(\tau = \infty) \geq \mathbb{P}(N(0, \infty) = \infty) = e^{-\int_0^\infty \lambda(\frac{\gamma}{s+1}) ds} \in (0, \infty).$$

□

3.4. Proof of Results in Section 2.4.

Proof of Proposition 23. Let us assume first that $\lambda(z) = \alpha z$. Then, we know that

$$(3.85) \quad \frac{N_t + \gamma}{(t+1)^\alpha} - \gamma = \int_0^t \frac{dM_s}{(s+1)^\alpha}$$

is a martingale. Therefore, for any $\epsilon > 0$, using $\mathbb{E}[N_t] = \gamma[(t+1)^\alpha - 1]$ from Proposition 21

$$(3.86) \quad \begin{aligned} \mathbb{P} \left(\sup_{0 \leq s \leq t} \left| \frac{N_s + \gamma}{(s+1)^\alpha} - \gamma \right| \geq \epsilon \gamma \right) &\leq \frac{\mathbb{E} \left[\left(\int_0^t \frac{dM_s}{(s+1)^\alpha} \right)^2 \right]}{\epsilon^2 \gamma^2} \\ &= \frac{\int_0^t \frac{\mathbb{E}[\lambda_s]}{(s+1)^\alpha} ds}{\epsilon^2 \gamma^2} \\ &= \frac{\int_0^t \frac{\alpha \gamma}{s+1} ds}{\epsilon^2 \gamma^2} \\ &\rightarrow 0, \end{aligned}$$

as $\gamma \rightarrow \infty$. Hence,

$$(3.87) \quad \begin{aligned} \mathbb{P} \left(\sup_{0 \leq s \leq t} \left| \frac{N_s}{\gamma} - [(s+1)^\alpha - 1] \right| \geq \epsilon \right) &= \mathbb{P} \left(\sup_{0 \leq s \leq t} \left| \frac{N_s + \gamma}{(s+1)^\alpha} - \gamma \right| (s+1)^\alpha \geq \epsilon \gamma \right) \\ &\leq \mathbb{P} \left(\sup_{0 \leq s \leq t} \left| \frac{N_s + \gamma}{(s+1)^\alpha} - \gamma \right| \geq \frac{\epsilon \gamma}{(t+1)^\alpha} \right) \\ &\rightarrow 0, \end{aligned}$$

as $\gamma \rightarrow \infty$. If $\lim_{z \rightarrow \infty} \frac{\lambda(z)}{z} = \alpha$, then for any $\delta > 0$, there exists K so that for any $z \geq K$, $(\alpha - \delta)z \leq \lambda(z) \leq (\alpha + \delta)z$. Uniformly for $0 \leq s \leq t$, $\frac{N_s + \gamma}{s+1} \geq \frac{\gamma}{t+1} \geq K$ for any $\gamma \geq K(t+1)$. Now using the results for $\lambda(z) = (\alpha + \delta)z$ and $\lambda(z) = (\alpha - \delta)z$ and let $\delta \rightarrow 0$, we proved Proposition 23. □

Proof of Proposition 24. For any $\epsilon > 0$ and fixed $t > 0$, for sufficiently large γ , $(\alpha - \epsilon)z^\beta \leq \lambda(z) \leq (\alpha + \epsilon)z^\beta$ for any $z \geq \frac{\gamma}{t+1}$. Since it holds for any $\epsilon > 0$, to prove Proposition 24, it suffices to consider the case $\lambda(z) = \alpha z^\beta$. Without loss of generality, let us take $\alpha = 1$. Let us use the Poisson embedding. Let $N^{(0)}$ be the Poisson process with intensity $\lambda(\frac{\gamma}{t+1}) = \frac{\gamma^\beta}{(t+1)^\beta}$ and the compensator

$$(3.88) \quad \int_0^t \lambda \left(\frac{\gamma}{s+1} \right) ds = \int_0^t \frac{\gamma^\beta}{(s+1)^\beta} ds = \frac{\gamma^\beta}{1-\beta} [(t+1)^{1-\beta} - 1].$$

It is easy to show that

$$(3.89) \quad \sup_{0 \leq s \leq t} \left| \frac{N_s^{(0)}}{\gamma^\beta} - \frac{1}{1-\beta} [(s+1)^{1-\beta} - 1] \right| \rightarrow 0,$$

in probability as $\gamma \rightarrow \infty$.

Conditional on $N^{(0)}$, let $N^{(1)}$ be the inhomogeneous Poisson process with intensity

$$(3.90) \quad \lambda \left(\frac{N_{t-}^{(0)} + \gamma}{t+1} \right) - \lambda \left(\frac{\gamma}{t+1} \right),$$

at time t . Inductively, conditional on $N^{(0)}, N^{(1)}, \dots, N^{(k)}$, $N^{(k+1)}$ is an inhomogeneous Poisson process with intensity

$$(3.91) \quad \lambda \left(\frac{N_{t-}^{(0)} + N_{t-}^{(1)} + \dots + N_{t-}^{(k)} + \gamma}{t+1} \right) - \lambda \left(\frac{N_{t-}^{(0)} + N_{t-}^{(1)} + \dots + N_{t-}^{(k-1)} + \gamma}{t+1} \right),$$

at time t . By the mean value theorem, and the assumption $0 < \beta < 1$,

$$(3.92) \quad \begin{aligned} & \lambda \left(\frac{N_{s-}^{(0)} + N_{s-}^{(1)} + \dots + N_{s-}^{(k)} + \gamma}{s+1} \right) - \lambda \left(\frac{N_{s-}^{(0)} + N_{s-}^{(1)} + \dots + N_{s-}^{(k-1)} + \gamma}{s+1} \right) \\ & \leq \beta \left(\frac{\gamma}{s+1} \right)^{\beta-1} N_s^{(0)} \\ & \leq \beta \left(\frac{\gamma}{t+1} \right)^{\beta-1} N_t^{(0)}. \end{aligned}$$

Therefore, by induction,

$$(3.93) \quad \begin{aligned} \mathbb{E}[N_t^{(k+1)}] & \leq \beta t \left(\frac{\gamma}{t+1} \right)^{\beta-1} \mathbb{E}[N_t^{(k)}] \\ & \leq \left(\beta t \left(\frac{\gamma}{t+1} \right)^{\beta-1} \right)^{k+1} \mathbb{E}[N_t^{(0)}]. \end{aligned}$$

For fixed t , for sufficiently large γ , we have $t \left(\frac{\gamma}{t+1} \right)^{\beta-1} \leq 1$. Hence, $\mathbb{E}[N_t^{(k+1)}] \leq \beta^{k+1} \mathbb{E}[N_t^{(0)}]$ and $\sum_{k=0}^{\infty} \mathbb{E}[N_t^{(k)}] \leq \frac{1}{1-\beta} \mathbb{E}[N_t^{(0)}] < \infty$. $N_t = \sum_{k=0}^{\infty} N_t^{(k)}$ is well defined and coincides with the self-exciting point process in our model from Poisson embedding. Moreover,

$$(3.94) \quad \begin{aligned} \mathbb{E} \left[\sum_{k=1}^{\infty} N_t^{(k)} \right] & \leq \sum_{k=1}^{\infty} \left(\beta \left(\frac{\gamma}{t+1} \right)^{\beta-1} \right)^k \mathbb{E}[N_t^{(0)}] \\ & \leq \beta \left(\frac{\gamma}{t+1} \right)^{\beta-1} \frac{\gamma^\beta}{1-\beta} [(t+1)^{1-\beta} - 1]. \end{aligned}$$

Therefore, we conclude that $\frac{\sum_{k=1}^{\infty} N_t^{(k)}}{\gamma^\beta} \rightarrow 0$ in probability as $\gamma \rightarrow \infty$. Hence, we proved the desired result. \square

3.5. Proofs of Results in Section 2.5.

Proof of Theorem 25. (i) The identity (2.34) holds from the definition of our model. The integrand is the infinitesimal probability that there are precisely k jumps on the time interval $[0, t]$ that occurs at $0 < t_1 < t_2 < \dots < t_k$.

(ii) This is a direct consequence of (i).

(iii) When $\lambda(z) = \alpha z$,

(3.95)

$$\begin{aligned}
\mathbb{P}(N_t = k) &= \int \cdots \int_{0 < t_1 < t_2 < \cdots < t_k < t} \alpha^k \prod_{j=1}^k \frac{\gamma + j - 1}{t_j + 1} \\
&\quad \cdot e^{-\int_0^{t_1} \frac{\alpha \gamma}{s+1} ds - \int_{t_1}^{t_2} \frac{\alpha(\gamma+1)}{s+1} ds - \cdots - \int_{t_k}^t \frac{\alpha(\gamma+k)}{s+1} ds} dt_1 dt_2 \cdots dt_k \\
&= \alpha^k \gamma(\gamma+1) \cdots (\gamma+k-1) \int \cdots \int_{0 < t_1 < t_2 < \cdots < t_k < t} \prod_{j=1}^k \frac{1}{t_j + 1} \\
&\quad \exp \left\{ -\alpha \gamma \log(t_1 + 1) + \alpha(\gamma+1) \log(t_1 + 1) - \alpha(\gamma+1) \log(t_2 + 1) \right. \\
&\quad \left. \cdots + \alpha(\gamma+k) \log(t_k + 1) - \alpha(\gamma+k) \log(t + 1) \right\} dt_1 \cdots dt_k \\
&= \alpha^k \gamma(\gamma+1) \cdots (\gamma+k-1) \frac{1}{(t+1)^{\alpha(\gamma+k)}} \\
&\quad \int \cdots \int_{0 < t_1 < t_2 < \cdots < t_k < t} \prod_{j=1}^k \frac{1}{(t_j + 1)^{1-\alpha}} dt_1 \cdots dt_k \\
&= \frac{1}{k!} \frac{\gamma(\gamma+1) \cdots (\gamma+k-1)}{(t+1)^{\alpha(\gamma+k)}} [(t+1)^\alpha - 1]^k \\
&= \binom{k+\gamma-1}{k} \left(1 - \frac{1}{(t+1)^\alpha} \right)^k \left(\frac{1}{(t+1)^\alpha} \right)^\gamma.
\end{aligned}$$

In other words, N_t follows a negative binomial distribution. Similarly,

(3.96) $\mathbb{P}(N_t = k + m | N_s = m)$

$$\begin{aligned}
&= \int \cdots \int_{s < t_1 < t_2 < \cdots < t_k < t} \alpha^k \prod_{j=1}^k \frac{\gamma + m + j - 1}{t_j + 1} \\
&\quad \cdot e^{-\int_s^{t_1} \frac{\alpha(\gamma+m)}{s+1} ds - \int_{t_1}^{t_2} \frac{\alpha(\gamma+m+1)}{s+1} ds - \cdots - \int_{t_k}^t \frac{\alpha(\gamma+m+k)}{s+1} ds} dt_1 dt_2 \cdots dt_k \\
&= \binom{k+m+\gamma-1}{k} \left(1 - \left(\frac{s+1}{t+1} \right)^\alpha \right)^k \left(\left(\frac{s+1}{t+1} \right)^\alpha \right)^{\gamma+m}.
\end{aligned}$$

□

Proof of Theorem 26. For any $\epsilon > 0$, there exists a constant $M(\epsilon)$ so that for any $z \geq M(\epsilon)$, $(\alpha - \epsilon)z \leq \lambda(z) \leq (\alpha + \epsilon)z$. Therefore, there exists some constant C_1 and C_2 that depend on ϵ , γ and t so that for any k

$$(3.97) \quad (\alpha - \epsilon)^k C_1 \leq \prod_{j=1}^k \frac{\lambda\left(\frac{\gamma+j-1}{t_j+1}\right)}{\frac{\gamma+j-1}{t_j+1}} \leq (\alpha + \epsilon)^k C_2.$$

And there also exist some C_3 and C_4 that may depend on ϵ , γ and t so that for any $0 < t_1 < \dots < t_k < t$,

$$\begin{aligned}
(3.98) \quad & -C_3 - \int_0^{t_1} \frac{(\alpha + \epsilon)\gamma}{s+1} ds - \int_{t_1}^{t_2} \frac{(\alpha + \epsilon)(\gamma + 1)}{s+1} ds - \dots - \int_{t_k}^t \frac{(\alpha + \epsilon)(\gamma + k)}{s+1} ds \\
& \leq - \int_0^{t_1} \lambda \left(\frac{\gamma}{s+1} \right) ds - \int_{t_1}^{t_2} \lambda \left(\frac{\gamma + 1}{s+1} \right) ds - \dots - \int_{t_k}^t \lambda \left(\frac{\gamma + k}{s+1} \right) ds \\
& \leq C_4 - \int_0^{t_1} \frac{(\alpha - \epsilon)\gamma}{s+1} ds - \int_{t_1}^{t_2} \frac{(\alpha - \epsilon)(\gamma + 1)}{s+1} ds - \dots - \int_{t_k}^t \frac{(\alpha - \epsilon)(\gamma + k)}{s+1} ds.
\end{aligned}$$

Hence, from the proof of Theorem 25, we have

$$\begin{aligned}
(3.99) \quad & C_1 e^{-C_3} \left(\frac{\alpha - \epsilon}{\alpha + \epsilon} \right)^k \binom{k + \gamma - 1}{k} \left(1 - \frac{1}{(t+1)^{\alpha + \epsilon}} \right)^k \left(\frac{1}{(t+1)^{\alpha + \epsilon}} \right)^\gamma \\
& \leq \mathbb{P}(N_t = k) \leq C_2 e^{C_4} \left(\frac{\alpha + \epsilon}{\alpha - \epsilon} \right)^k \binom{k + \gamma - 1}{k} \left(1 - \frac{1}{(t+1)^{\alpha - \epsilon}} \right)^k \left(\frac{1}{(t+1)^{\alpha - \epsilon}} \right)^\gamma.
\end{aligned}$$

Since it holds for any $\epsilon > 0$, we proved (2.38). \square

Proof of Theorem 27. The results for the case $\lim_{z \rightarrow \infty} \frac{\lambda(z)}{z^\beta} = \alpha$ can be reduced to the case $\lambda(z) = \alpha z^\beta$ by following the similar arguments as in the proof of Theorem 26. So let us assume that $\lambda(z) = \alpha z^\beta$.

$$\begin{aligned}
(3.100) \quad & \mathbb{P}(N_t = k) = \int \dots \int_{0 < t_1 < t_2 < \dots < t_k < t} \alpha^k \prod_{j=1}^k \left(\frac{\gamma + j - 1}{t_j + 1} \right)^\beta \\
& \quad \cdot e^{-\int_0^{t_1} \frac{\alpha \gamma^\beta}{(s+1)^\beta} ds - \int_{t_1}^{t_2} \frac{\alpha(\gamma+1)^\beta}{(s+1)^\beta} ds - \dots - \int_{t_k}^t \frac{\alpha(\gamma+k)^\beta}{(s+1)^\beta} ds} dt_1 dt_2 \dots dt_k \\
& \leq \int \dots \int_{0 < t_1 < t_2 < \dots < t_k < t} \alpha^k \prod_{j=1}^k (\gamma + j - 1)^\beta dt_1 dt_2 \dots dt_k \\
& = \alpha^k t^k \frac{1}{k!} \prod_{j=1}^k (\gamma + j - 1)^\beta.
\end{aligned}$$

Therefore,

$$(3.101) \quad \limsup_{\ell \rightarrow \infty} \frac{1}{\ell \log \ell} \log \mathbb{P}(N_t \geq \ell) \leq -(1 - \beta).$$

On the other hand,

(3.102)

$$\begin{aligned} \mathbb{P}(N_t = k) &= \int \cdots \int_{0 < t_1 < t_2 < \cdots < t_k < t} \alpha^k \prod_{j=1}^k \left(\frac{\gamma + j - 1}{t_j + 1} \right)^\beta \\ &\quad \cdot e^{-\int_0^{t_1} \frac{\alpha \gamma^\beta}{(s+1)^\beta} ds - \int_{t_1}^{t_2} \frac{\alpha(\gamma+1)^\beta}{(s+1)^\beta} ds - \cdots - \int_{t_k}^t \frac{\alpha(\gamma+k)^\beta}{(s+1)^\beta} ds} dt_1 dt_2 \cdots dt_k \\ &\geq \alpha^k t^k \frac{1}{k!} \prod_{j=1}^k \left(\frac{\gamma + j - 1}{t + 1} \right)^\beta e^{-\alpha(\gamma+k)^\beta t}. \end{aligned}$$

Therefore,

$$(3.103) \quad \liminf_{\ell \rightarrow \infty} \frac{1}{\ell \log \ell} \log \mathbb{P}(N_t \geq \ell) \geq -(1 - \beta),$$

and we proved the desired result. \square

4. CONCLUSION AND OPEN PROBLEMS

In this paper, we studied a class of self-exciting point processes. We proved that the limit in the law of large numbers is a fixed point of the rate function. When the rate function is linear, explicit formulas were obtained for the mean, variance and covariance. Central limit theorem and large deviations were also studied. Finally, for a fixed time interval, we obtain the asymptotics for the tail probabilities. Here is a list of open problems that are interesting to investigate in the future.

- When there are more than one fixed point of $x = \lambda(x)$, say there are exactly two stable fixed points $x_1 < x_2$, we made a plot of the probability p_1 and p_2 that the process $\frac{N_t}{t}$ converges to x_1 and x_2 respectively as a function of the initial condition γ . From Figure 5, the simulations suggest that p_1, p_2 are monotonic in γ . Is that always true? Can we compute $p_1(\gamma), p_2(\gamma)$ analytically or at least obtain asymptotics for $\gamma \rightarrow 0^+$ and $\gamma \rightarrow \infty$?
- So far, we have concentrated on the case when $x = \lambda(x)$ has finitely many fixed points. It is natural to ask what if there are infinitely many fixed points, or more precisely, what if the Lebesgue measure of the set of fixed points is positive, then, what will be the limiting distribution of $\frac{N_t}{t}$ like as $t \rightarrow \infty$?. In Figure 6, we consider a piecewise $\lambda(x)$ that coincides with x on the interval $[2, 3]$ and $[4.5, 5]$ and Figure 7 illustrates the limiting set of $\frac{N_t}{t}$ as time $t \rightarrow \infty$. Figure 7 suggests that the limiting set is supported on $[2, 3]$ and $[4.5, 5]$.
- Sometimes, a fixed point of $x = \lambda(x)$ can be neither stable or unstable. It is possible to have a saddle point, i.e., stable from one side and unstable from the other. Figure 8 gives such an example in which $\lambda(x)$ is piecewise linear and there is a stable fixed point at $x = 5$ and two saddle points at $x = 2$ and $x = 6.5$. Can we analyze this situation?
- Can we relax the assumption $\lambda(\cdot) \leq C_0 < \infty$ in Theorem 7 for the large deviations? Can this assumption be relaxed to $\lim_{x \rightarrow \infty} \frac{\lambda(x)}{x} = 0$?
- We obtained explicit formulas for the mean, variance and covariance of N_t when $\lambda(x)$ is linear (here $x = \lambda(x)$ may not have a fixed point). Can we at least obtain the asymptotics for the mean, variance and covariance for large t when $\lambda(x)$ is nonlinear?

- We can also consider a d -dimensional simple point process $(N_t^{(1)}, \dots, N_t^{(d)})$, where $N_t^{(i)}$ has intensity at time t given by

$$\lambda_t^{(i)} = \frac{\sum_{j \neq i} a_{ij} N_{t-}^{(j)}}{t+1} + b_i.$$

More generally, we can consider for example $\lambda_t^{(i)} = \lambda_i(\frac{1}{t+1} \sum_{j \neq i} a_{ij} N_{t-}^{(j)})$ for nonlinear $\lambda_i(\cdot)$. The d -dimensional process $(N_t^{(1)}, \dots, N_t^{(d)})$ is thus mutually exciting. Can we do the similar analysis to study the d -dimensional process as in our paper?

ACKNOWLEDGEMENTS

The authors are grateful to Maury Bramson and Wenqing Hu for helpful discussions.

REFERENCES

- [1] Bacry, E., S. Delattre, M. Hoffmann and J. F. Muzy. (2013). Scaling limits for Hawkes processes and application to financial statistics. *Stochastic Processes and their Applications* **123**, 2475-2499.
- [2] Bacry, E., S. Delattre, M. Hoffmann and J. F. Muzy. (2011). Modeling microstructure noise with mutually exciting point processes. Preprint.
- [3] Blundell, C., Heller, K. A. and J. M. Beck. Modelling reciprocating relationships with Hawkes processes. Preprint, 2012.
- [4] Bordenave, C. and Torrisi, G. L. (2007). Large deviations of Poisson cluster processes. *Stochastic Models*, **23**, 593-625.
- [5] Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Probab.* **24**, 1563-1588.
- [6] Chavez-Demoulin, V., Davison, A. C. and A. J. McNeil. (2005). Estimating value-at-risk: a point process approach. *Quantitative Finance*. **5**, 227-234.
- [7] Chornoboy, E. S., Schramm, L. P. and A. F. Karr. (1988). Maximum likelihood identification of neural point process systems. *Biol. Cybern.* **59**, 265-275.
- [8] Cox, J. T. and G. Grimmett. (1984). Central limit theorems for associated random variables and the percolation model. *Annals of Probability*. **12**, 514-528.
- [9] Crane, R. and D. Sornette. (2008) Robust dynamic classes revealed by measuring the response function of a social system. *Proc. Nat. Acad. Sci. USA* **105**, 15649.
- [10] Dembo, A. and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd Edition, Springer, 1998
- [11] Errais, E., Giesecke, K. and Goldberg, L. (2010). Affine point processes and portfolio credit risk. *SIAM J. Financial Math.* **1**, 642-665.
- [12] Feller, W., *An Introduction to Probability Theory and Its Applications*, Volume I and Volume II, 2nd Edition, New York, 1971.
- [13] Fierro, R., Leiva, V. and J. Møller. The Hawkes process with different exciting functions and its asymptotic behavior. Preprint.
- [14] Foo, J. and Leder, K. (2013). Dynamics of cancer recurrence. *Annals of Applied Probability* **23**, 1437-1468.
- [15] Foo, J., Leder, K., and Zhu, J. (2014). Escape times for branching processes with random mutational fitness effects. *Stochastic Processes and their Applications*. **124**, 3661-3697.
- [16] Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58**, 83-90.
- [17] Hawkes, A. G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *J. Appl. Prob.* **11**, 493-503.
- [18] Hawkes, A. G. and L. Adamopoulos. (1973). Cluster models for earthquakes-regional comparisons. *Bull. Int. Statist. Inst.* **45**, 454-461.
- [19] Karabash, D. and Zhu, L. Limit theorems for marked Hawkes processes with application to a risk model. Preprint.

- [20] Kwieciński, A. and R. Szekli. (1996). Some monotonicity and dependence properties of self-exciting point processes. *Annals of Applied Probability*. **6**, 1211-1231.
- [21] Lipster, R. S. and A. N. Shiryaev, *Statistics of Random Processes II. Applications*, 2nd Edition, Springer, 2001.
- [22] Merhdad, B. and L. Zhu. On the Hawkes process with different exciting functions. Preprint.
- [23] Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *J. Amer. Statist. Assoc.* **83**, 9-27.
- [24] Pernice, V., Staude B., Carndanobile, S. and S. Rotter. (2012). How structure determines correlations in neuronal networks. *PLoS Computational Biology*. **85**:031916.
- [25] Pernice, V., Staude B., Carndanobile, S. and S. Rotter. (2011). Recurrent interactions in spiking networks with arbitrary topology. *Physical Review E*. **7**:e1002059.
- [26] Varadhan, S. R. S., *Large deviations*, The Annals of Probability, Vol. 36, No. 2, 397-419, 2008
- [27] Varadhan, S. R. S., *Large Deviations and Applications*, SIAM, 1984
- [28] Zhu, L. Limit theorems for a Cox-Ingersoll-Ross process with Hawkes jumps. (2014). *Journal of Applied Probability*. **51**, 699-712.
- [29] Zhu, L. (2013). Moderate deviations for Hawkes processes. *Statistics & Probability Letters*. **83**, 885-890.
- [30] Zhu, L. Large deviations for Markovian nonlinear Hawkes processes. *to appear in Annals of Applied Probability*.
- [31] Zhu, L. Process-level large deviations for nonlinear Hawkes point processes. (2014). *Annales de l'Institut Henri Poincaré*. **50**, 845-871.
- [32] Zhu, L. Central limit theorem for nonlinear Hawkes processes. (2013). *Journal of Applied Probability*. **50**, 760-771.

SCHOOL OF MATHEMATICS
 UNIVERSITY OF MINNESOTA-TWIN CITIES
 206 CHURCH STREET S.E.
 MINNEAPOLIS, MN-55455
 UNITED STATES OF AMERICA
 E-mail address: yangx953@umn.edu, zhul@umn.edu

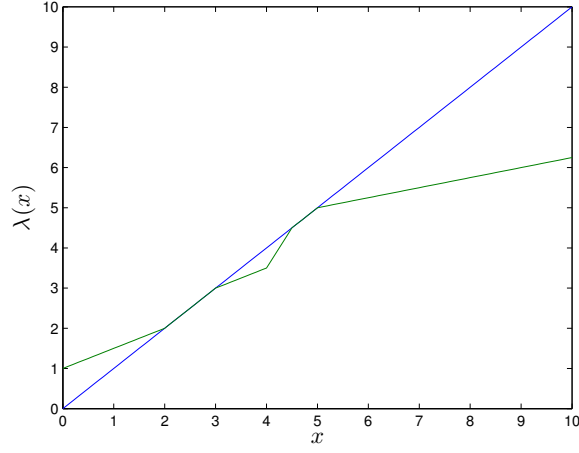


FIGURE 6. We consider a piecewise function $\lambda(x)$ defined as $1 + 0.5x$ for $x < 2$, x for $2 \leq x < 3$, $1.5 + 0.5x$ for $3 \leq x < 4$, $-4.5 + 2x$ for $4 \leq x < 4.5$, x for $4.5 \leq x < 5$ and $3.75 + 0.25x$ for $x \geq 5$. The set of the fixed points of $x = \lambda(x)$ is $[2, 3] \cup [4.5, 5]$.

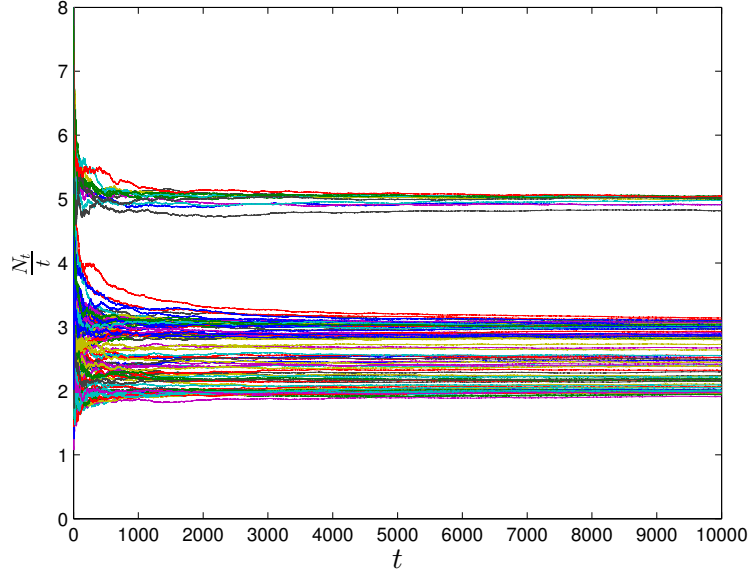


FIGURE 7. We choose the initial starting point as 3.5. The function $\lambda(x)$ is defined in Figure 6. We simulate 100 sample paths and the illustration suggests that the limiting set of $\frac{N_t}{t}$ as time $t \rightarrow \infty$ is supported on $[2, 3]$ and $[4.5, 5]$.

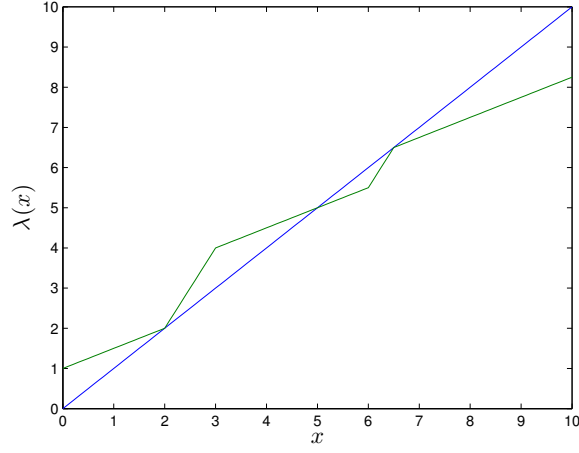


FIGURE 8. We consider a piecewise function $\lambda(x)$ defined as $1 + 0.5x$ for $x < 2$, $-2 + 2x$ for $2 \leq x < 3$, $2.5 + 0.5x$ for $3 \leq x < 6$, $-6.5 + 2x$ for $6 \leq x < 6.5$, and $3.25 + 0.5x$ for $x \geq 6.5$. The set of the fixed points consist of a stable fixed point at $x = 5$ and two saddle points at $x = 2$ and $x = 6.5$.

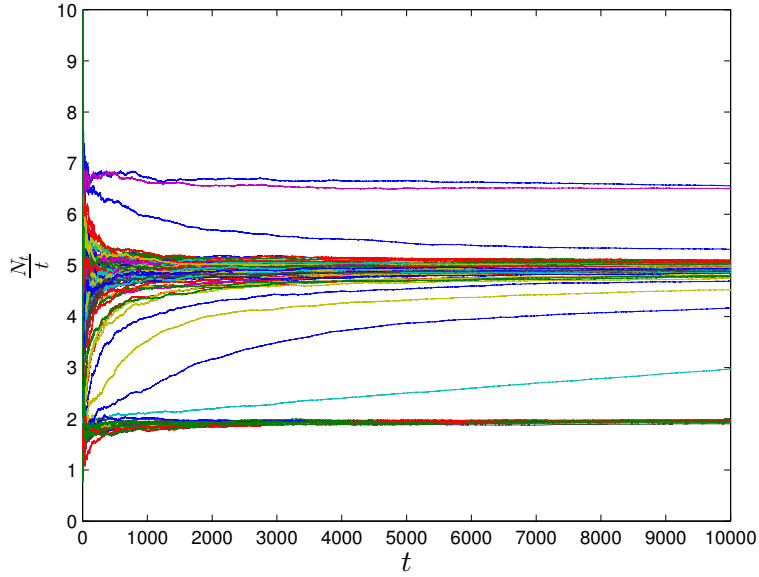


FIGURE 9. We choose the initial starting point as 2.5. The function $\lambda(x)$ is defined in Figure 8. We simulate 100 sample paths and the illustration suggests that the limiting set of $\frac{N_t}{t}$ as time $t \rightarrow \infty$ is supported on $\{2, 5, 6.5\}$.